MANGO MATURITY CLASSIFICATION USING MACHINE LEARNING

BY

XIANGCHEN DONG

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ENGINEERING IN ELECTRICAL AND COMPUTER ENGINEERING COLLEGE OF ENGINEERING

GRADUATE SCHOOL, RANGSIT UNIVERSITY

ACADEMIC YEAR 2024

Thesis entitled

MANGO MATURITY CLASSIFICATION USING MACHINE LEARNING

by

XIANGCHEN DONG

was submitted in partial fulfillment of the requirements
for the degree of Master of Engineering in Electrical and Computer Engineering

Rangsit University
Academic Year 2024

Assoc.Prof. Opas Chutatape, Ph.D.
Examination Committee Chairperson

Asst.Prof. Supattana Nirukkanaporn, D.Eng.
Member

Assoc.Prof. Rong Phoophuangpairoj, Ph.D.
Member and Advisor

Approved by Graduate School

(Prof. Suejit Pechprasarn, Ph.D.)
Dean of Graduate School
December 30, 2024

Acknowledgements

I would like to extend my sincere gratitude to my advisor, Assoc. Prof. Rong Phoophuangpairoj, Ph.D., for his invaluable guidance and profound insights throughout my research journey, particularly in selecting and exploring the research topic. His mentorship has been instrumental in shaping my academic pursuits and enhancing my understanding of this field.

Xiangchen Dong
Researcher

6509121	:	Xiangchen Dong
Thesis Title	:	Mango Maturity Classification Using Machine Learning
Program	:	Master of Engineering in Electrical and Computer Engineering
Thesis Advisor	:	Assoc.Prof. Rong Phoophuangpairoj, Ph.D.

Abstracts

Mangoes, predominantly cultivated in tropical and subtropical regions, are cherished for their sweet and sour taste, pleasant aroma, and rich vitamin content. This study focused on classifying mango ripeness using various machine learning classifiers. Images of mangoes at different ripeness stages were collected and used to train classifiers, including Convolutional Neural Network (CNN), MobileNet, ResNet50, and VGG16. The experimental results demonstrated that CNN, MobileNet, ResNet50, and VGG16 achieved accuracies of 81.30%, 85.11%, 73.66%, and 90.08%, respectively. VGG16 attained the highest classification accuracy, with classification accuracies from class 0 to class 5 being 98.85%, 98.85%, 95.80%, 95.80%, 95.42%, and 95.42%, respectively.

Building on these results, the second part of the study utilized the top-performing VGG16 model to output softmax layer data from mango images. This softmax data, combined with the mangoes' RGB and Lab color data, was used to make predictions via linear regression and artificial neural networks (ANNs). The comparative analysis revealed that ANN generally yielded better predictive performance.

(Total 88 pages)

Keywords: Machine Learning, Mango, CNN, VGG16, ResNet50, ANN, Brix, pH

Student's Signature …... Thesis Advisor's Signature ....

Table of Contents

Table of Contents (continued)

Table of Contents (continued)

List of Tables

List of Tables (continued)

Page

Tables

List of Tables (continued)

List of Figures

List of Figures (continued)

Page

Figures

Chapter 1

Introduction

1.1 Background and Significance of the Problem

Mangoes are one of the most popular and tasty fruits grown and consumed worldwide. Thailand is one of the world's largest mango producers and exporters (Chomchalow & Songkhla, 2008). Mangoes can be classified according to their maturity levels. But the majority of previous research only focused on the differentiation between ripe and unripe fruits (Rizzo, Marcuzzo, Zangari, Gasparetto, &Albarelli, 2023). In embarking on the journey of mango maturity classification through machine learning, (Worasawate et al., 2022) developed machine learning models for predicting the ripeness stage of mangoes at harvest based on weight, skin color, capacitance of each mango, weight/capacitance ratio. (Shahriar et al., 2023) studied the methods for classifying five different types of mangoes using the transfer learning (TL) approach. The research also delved into the realm of CNNs (Wu, 2017), investigating the possibilities of well-known architectures such as ResNet (Çınar et al., 2021), VGG16 (Rezende, Ruppert, Carvalho, Theophilo, Ramos, &Geus, 2018), and MobileNet (Nan, Ju, Hua, Zhang, &Wang, 2022).

Ripeness assessment of mangoes plays a key role in consumer choice, export, and food industries. Predicting the ripeness stage of mangoes (Eyarkai Nambi, Thangavel, Manickavasagan, &Shahir, 2017) would also aid in decision-making during transportation based on the time necessary.

1.1.1 Importance of fruit ripeness assessment

Fruit ripeness assessment is a key factor in both agricultural practices and the food industry for several compelling reasons (Seymour et al., 2013). Accurate assessment allows growers and producers to harvest fruit just as it reaches peak ripeness, maximizing both yield and quality. This optimal timing ensures superior taste, better nutritional value, and longer shelf life, which together improve the overall quality of the product.

It is simpler to choose appropriate storage conditions and transportation schedules when one has a better understanding of fruit maturity. This information efficiently minimizes post-harvest losses, keeps fruit fresh, and stops premature rotting or decay during storage and transit (Strano et al., 2022).

In addition, by assessing the degree of ripeness of fruits, an adjustment of market supply to consumer demand can be achieved. Producers can adapt harvest and distribution plans to market demand and ensure a steady supply of ripe fruit (El-Ramady, Domokos-Szabolcsy, Abdalla, Taha, &Fάri,2015). This approach greatly increased both market competitiveness and customer satisfaction. Reducing over- or under-ripeness is one of the main ways reduces food waste. Accurately scheduling the collection and delivery of fruit lowers the quantity of unsold or rejected fruit, which lowers waste.

From an economic perspective, longer shelf lives, improved product quality, and decreased waste are advantageous to both consumers and supply chain stakeholders (Ciccullo et al., 2021). Improved inventory management techniques result in higher revenue and cost savings from an efficient maturity assessment.

Fruit maturity assessment fundamentally plays an essential role in improving agricultural productivity, ensuring food quality and safety, minimizing waste, meeting market demands, and promoting the financial sustainability of the entire fruit supply chain (Vidanapathirana, 2019).

1.1.2 Overview of mango ripeness assessment based on skin color

Assessment of mango ripeness through skin color analysis is an important non-destructive method widely used in agricultural practices and the food industry (Zhena, Hashima, &Maringgala , 2020). This assessment relied on visible changes in mango skin color, which serve as a visual cue indicating the fruit's stage of ripeness.

As a mango ripens, its peel undergoes various color transitions from green to bright yellow, orange or red (Prabhu & Mamatha, 2022). These skin color changes correlate with internal biochemical changes, thus providing a reliable visual indicator for assessing ripeness.

Analyzing changes in skin color, image processing techniques, and machine learning algorithms, specifically deep learning models, accurately measure and categorize mango ripeness (Naranjo-Torres et al., 2020). They examined machine learning methods to interpret color alterations, enabling the classification of mangoes into several phases of maturity.

The field of transfer learning was explored in the study. The goal was to optimize the adaptable VGG16 by fine-tuning it (Pardede et al., 2021) to enhance the accuracy and adaptability of the model, enabling farmers, distributors, and customers to utilize it as a dynamic tool. The anticipated benefits of the research extend far beyond academia. Imagine a future where precision optimizes harvesting, minimizes waste, supplies markets with ripe fruits at the right time, and boosts consumer satisfaction. The study aimed to usher in a new era in which agriculture and technology marry to revolutionize not only the way that mangoes were classified as ripe but also the way that we interact with fruits from the farm to the table.

The non-destructive method based on skin color analysis offers an efficient way to assess mango ripeness, thereby enabling informed decisions in harvesting, sorting, and distribution practices. Its reliability, efficiency, and non-invasive nature contribute significantly to optimizing the mango supply chain and ensuring market readiness.

1.2 Research Objectives

1.2.1 Develop machine learning models to classify mango ripeness from images.

Establish an efficient machine learning model capable of accurately classifying the ripeness of mangoes. This objective entails collecting a sufficient quantity and diversity of mango image datasets, preprocessing the data, selecting an appropriate machine learning architecture, and ensuring the model's accuracy and generalization through training and validation.

1.2.2 Explore the correlation between mango ripeness features and physicochemical properties

Utilize data analysis and statistical methods to investigate the correlation between mango features and pH value, sweetness, and basic color properties. This objective entails conducting laboratory tests on collected mango samples to obtain relevant data, as well as employing an appropriate analytical tool to uncover potential relationships.

1.2.3 Establish a predictive model for ripeness and physicochemical properties.

Based on the previous research findings, develop a predictive model capable of estimating the pH value and sweetness of mangoes based on their colors, appearance, and other characteristics. This objective requires integrating data analysis and deep learning techniques, potentially involving feature engineering, model optimization, and validation steps to ensure the accuracy and practicality of the model.

1.2.4 Validate the reliability and applicability of the model

Verifying the developed model's efficacy and dependability in practical applications is the goal. This involves collaboration with mango producers or processors to collect data from actual production environments and conduct field tests and validation of the model to assess its performance and application.

Chapter 2

Literature Review

The analysis of fruit ripeness plays a crucial role in both agricultural production and consumer satisfaction. With the advent of advanced technologies, particularly artificial intelligence and image processing, researchers have increasingly turned to CNNs for their efficacy in image classification tasks (Kangune, Kulkarni, & Kosamkar, 2019) were explored. Key topics included an overview of fruit ripeness analysis methodologies, the application of CNNs in image classification (Elngar et al., 2021), and the specific advancements in utilizing CNNs for fruit ripeness analysis (Singh et al., 2022). In addition, this review looked into the features and uses of several popular CNN architectures, like ResNet50, VGG16, and MobileNet, focusing on how well they work for figuring out how ripe a fruit is. The Artificial Neural Networks (ANN) and Support Vector Machines (SVM) (Kalantar, Pradhan, Naghibi, Motevalli, & Mansor, 2017), in data classification as well as the utilization of linear regression in data analysis, were also studied to classify mango ripeness and determine the properties of mangoes.

This review, by synthesizing existing literature in these areas, aimed to provide a comprehensive understanding of the current state-of-the-art methodologies and technologies employed in fruit ripeness analysis and classification.

2.1 Fruit Ripeness Assessment Techniques

Fruit ripeness assessment involves a number of techniques that are critical to determining the ripeness and quality of fruits. Traditional methods, which include visual inspection and manual squeezing, rely on observable changes such as changes in colors, firmness, and aroma to detect different stages of ripeness (Mkhathini, 2014).

Technological advancements have introduced non-destructive approaches to fruit ripeness assessment. Techniques such as spectral imaging enable detailed analysis of fruit characteristics based on spectral reflectance patterns (Yang et al., 2017). Near-infrared spectroscopy (NIRS) facilitated the rapid determination of internal characteristics, such as sugar content, without causing any damage to the fruit (Magwaza et al., 2011). In addition, ultrasonic methods evaluated internal properties such as pulp firmness and sugar content by analyzing sound wave responses or ultrasonic signals (Bureau, 2009).

Mature assessment is also still being revolutionized by new technologies. For instance, spectroscopy and imaging were combined in hyperspectral imaging to provide comprehensive data and enable precise fruit ripeness assessment (Lorente et al., 2011). Moazzem (2023) reported that the fruit's aroma and taste profile were revealed through the use of Electronic Nose (E-Nose) and Electronic Tongue (E-Tongue) systems for the detection of volatiles and taste compounds, respectively.

Researchers have investigated various methods for delimiting and categorizing different types of fruit in the field of fruit classification. Traditional approaches relied on conventional computer vision and image processing techniques and emphasized features such as color, shape, and texture in classification (Ngugi, Abelwahab, &Abo-Zahhad, 2021).

Studies employing fuzzy logic and MLP neural network classification techniques had successfully classified apples (Miriti, 2016), oranges, and bananas based on their visual characteristics.

Furthermore, researchers investigated the fusion of multimodal data sources, such as combining image data with sensor data like infrared spectra or sound, with the aim of enhancing the accuracy of fruit classification (Ignatious et al., 2023).

Researchers have integrated hyperspectral imaging with CNNs to distinguish between different grape varieties based on their spectral signatures (Nguyen et al., 2021).

Understanding the developments and limitations in these diverse areas contributed to the comprehensive background required for the present study on mango image classification using various convolutional neural networks. These technological advancements in fruit ripeness assessment contributed to rapid, non-invasive, and precise evaluation methods. They significantly enhanced harvest management and quality control practices across various fruit varieties and species, ensuring better decision-making processes and improved fruit quality throughout the supply chain.

2.2 CNNs in Image Classification

Classification is an important task in machine learning and consists of various algorithms to categorize data into classes or groups (Osisanwo, 2017). It plays an important role in many fields, including image recognition, natural language processing, medical diagnosis, and more. The ability to accurately analyze data is critical to decision-making and predictive modeling.

CNNs have emerged as a powerful tool for classification tasks, especially in image recognition. These networks performed well in hierarchical learning represented by raw data, eliminating the need for manual extraction (Yang et al., 2017). Their deep model included a convolutional algorithm that removed and learned complex features, followed by a convolutional algorithm that minimized the learned features, and a fully integrated fruit classification that evolved from traditional computer vision methods. It was found that CNNs , which use convolutional layers to pull out features, worked better than older methods like decision trees and support vector machines (Khatun et al., 2020).

According to studies, CNNs were proficient at classifying fruits. The integration of multimodal data, such as hyperspectral imaging with CNNs, improved accuracy in grape variety classification (Yang et al., 2021). Additionally, advancements in adversarial learning, data augmentation, and transfer learning contributed to robust fruit classification systems.

Real-time fruit classification challenged and applications, such as mango ripeness level classification, further underscored the relevance of CNNs in practical scenarios.

2.2.1 CNNs in fruit ripeness classification

The importance of CNNs in the classification of fruit ripeness stemed from their unique ability to handle image-based functions and change the measurement of fruit ripeness. Artificial neural networks highlight their advantages in classification (Mazen & Nashat, 2019).

CNNs are highly efficient at extracting features from images, making them powerful tools in fruit ripeness assessment. It is advised to use these networks since they effectively capture significant characteristics like color shifts, minute details, and subtly patterned fruit at various stages (Aherwadi et al., 2022). CNN-based methodologies automated the classification process, reducing the need for manual intervention (Wang et al., 2017). This model was capable of learning and identifying ripeness patterns autonomously from training data. They provided high precision and consistency in evaluating fruit ripeness in various datasets to ensure reliability and reproducibility.

CNNs exhibit scalability and the ability to learn relationships in image data. As the volume of data increases, CNNs can adapt and refine their prediction capabilities, improving their accuracy in classifying the tasks completed in time. This ability to learn allows for continuous improvement, making it suitable for changing the fruit ripeness measurements required.

In summary, the use of CNNs in fruit ripeness classification provides a good solution by extracting quality characteristics and improving the performance of the classification process. The CNNs greatly improve the precision and efficiency of the fruit ripeness assessment, thus revolutionizing the practices in agriculture and food industries.

2.2.2 ResNet, VGG16 and MobileNet

In the dynamic landscape of fruit classification, the advent of convolutional neural networks (CNNs) has been accompanied by the exploration of various architectures to enhance accuracy and efficiency. Traditional methods, including support vector machines and decision trees, have given way to sophisticated CNN architectures (Cervantes, Garcia-Lamont, Rodríguez-Mazahua,&Lopez, 2020). Among these, ResNet, VGG16, and MobileNet have emerged as prominent choices.

ResNet, characterized by residual blocks, addresses the vanishing gradient problem, enabling the training of deeper networks. VGG16, with its deep and homogeneous architecture, extensively utilizes 3x3 convolutional filters to enhance its classification performance. MobileNet, on the other hand, focuses on lightweight and efficient architectures suitable for mobile and edge devices.

Studies incorporating these architectures into fruit classification tasks demonstrated their efficacy. ResNet's ability to capture intricate features in fruit images, VGG16's success in learning hierarchical representations, and MobileNet's efficiency in resource-constrained environments showcased the versatility of these models (Alqahtani, 2024).

In addition to architectural considerations, methods such as transfer learning, fine-tuning, and ensemble learning are commonly employed to further enhance the performance of these CNNs in fruit classification tasks. The integration of these advanced architectures and methods contributes to the ongoing evolution of fruit classification techniques.

Understanding the strengths and limitations of ResNet, VGG16, and MobileNet in fruit classification is crucial for the present study on mango image classification. This knowledge aids in selecting the most suitable architecture and methodology for achieving accurate and efficient results.

Machine learning algorithms have proven highly effective in solving classification problems, especially image recognition tasks. They alleviate the need for manual feature extraction by autonomously learning pertinent features from input images, achieving commendable accuracy rates. In a study focusing on banana ripeness classification, a CNN demonstrated remarkable performance, achieving an average classification accuracy of 91.21% (Yamparala et al., 2020).

MobileNet, known for its ability to balance accuracy and computational speed, utilizes depthwise separable convolutions, replacing standard convolution filters. By employing depthwise convolution and pointwise convolution layers, MobileNet significantly reduced computation time while maintaining acceptable accuracy levels (Bouguezzi et al. , 2021).

ResNet50, a convolutional neural network architecture comprising 50 layers, and VGG16, another deep convolutional neural network, represent significant advancements in computer vision tasks such as image classification and object detection. VGG16, featuring 16 convolutional layers along with fully connected layers, stands out for its robust performance and design. Its versatility and superior performance make it widely applicable to various image classification tasks.

In comparative evaluations involving fruit and vegetable classification, the VGG16 model outperformed the ResNet50 technique slightly (Mimma et al., 2022). VGG16's architectural design and performance contributed to its superior performance in discerning various types of fruits and vegetables, showcasing its effectiveness in image classification tasks. These advancements and comparative analyses demonstrated the varied strengths and optimizations within different convolutional neural network architectures for image classification tasks, offering insights into their applications across diverse domains.

The ripening process of mangoes involves a notable transition in their skin color, shifting from green to yellow. This transformation serves as a prominent indicator, suggesting the feasibility of determining mango ripeness based on color changes (Supekar & Wakode, 2020). The changes in color correspond to variations in taste, with mangoes evolving from sour to sweet as they ripen. Each stage of ripening exhibits distinct taste characteristics, underscoring the importance of classifying mangoes based on their ripening stages (Prabhu et al., 2024).

Image processing techniques coupled with machine learning methodologies have proven effective in assessing mango ripeness, size, and shape. This method efficiently categorizes mangoes into different stages of ripeness, facilitating quality assessment and market readiness.

The VGG network, known for its robustness in image classification tasks, demonstrated an exceptional accuracy in classifying various types of images. The VGG network showed obvious advantages in terms of accuracy and precision, highlighting its effectiveness on various types of products (Yang & Xu, 2021).

The ripening stages of Kaew Kamin mangoes, a popular variety, follow a progression from green to yellow, which is accompanied by distinct flavor transitions. These stages can be categorized into six phases based on skin color, softness, and hardness. These six stages described the development of mangoes from their initial green state to a ripe, yellow stage (see Figure 2.1), with each stage having its own unique flavor profile.

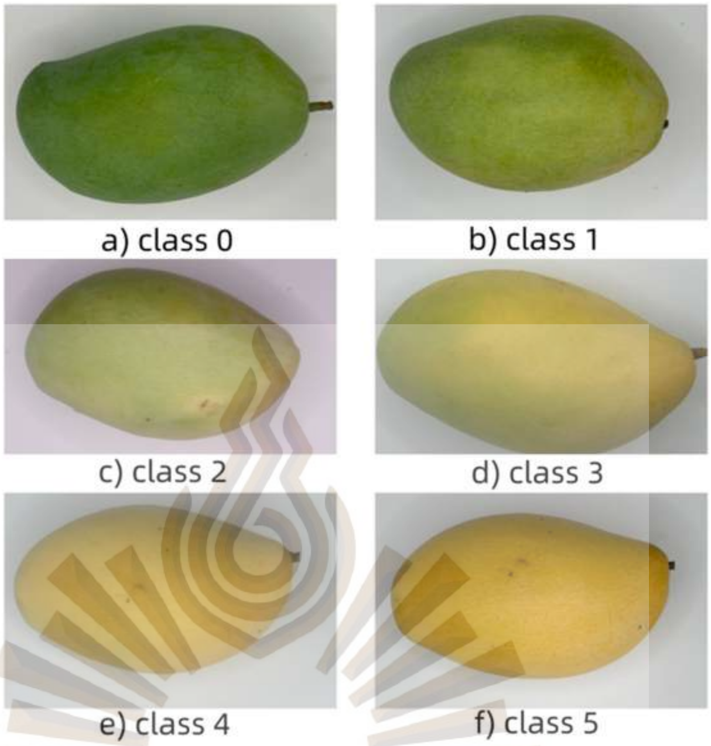

Figure 2.1 Classification of mango ripening stages.

Source: Researcher, 2024

Mangoes could be classified into six different ripeness levels based on their color and texture (Mim et al., 2018). Class 0 mangoes are dark green with a harder texture. Class 1 mangoes are light green and also have a hard texture. As mangoes begin to ripen, a small amount of yellow appears in the green skin, which is classified as Class 2, and they still have a hard texture. Class 3 mangoes mostly appear yellow and have a slightly softer texture. By Class 4, mangoes turn completely light yellow and have a softer texture. Finally, Class 5 mangoes exhibit a deep yellow color, have a very soft texture, and have a rich taste.

This comprehensive understanding of mango ripening stages, coupled with advanced image processing and machine learning techniques, provides an efficient framework for accurately assessing mango ripeness, enabling precise classification and quality assessment throughout the entire mango ripening process.

2.3 Overview of ANN

ANN is a computational model that follows the structure of human neural networks, which consist of many layers of neurons. ANN can model and predict nonlinear data. In the prediction of mango maturity and nutritional content, ANN can learn the relationship between images and data, identify mangoes in different stages, and predict changes in the food. An ANN usually consists of an input layer, a hidden layer, and an output layer. By adjusting parameters such as the number of layers, nodes, and network functions, the model can be optimized and the prediction performance can be improved. Artificial neural networks are effective in capturing characteristics and patterns and have accuracy and precision in predicting mango ripeness and its nutritional values.

Chapter 3

Research Methodology

3.1 Materials and Methods

3.1.1 Data for creating image classification

The study utilized Kaew Kamin mangoes (scientific name: Mangifera indica L.) for training and testing purposes. The image dataset was divided into two sets: an 80-20% split and a 70-30% split. The training set comprised 1080 mango images, with class distributions of 160, 130, 230, 280, 140, and 140 images for classes 0, 1, 2, 3, 4, and 5, respectively. The testing set consisted of 262 images, with class distributions of 38, 26, 64, 70, 32, and 32 images for classes 0, 1, 2, 3, 4, and 5, respectively. The second split used 940 training images (140, 120, 190, 240, 120, and 130 for classes 0 to 5) and 402 testing images (58, 34, 104, 110, 52, and 42 for classes 0 to 5). Class 0 and class 1 were relatively rare, due to the fact that Class 2 and Class 3 are being harvested and sold in the market, so there are more images for these two levels.

3.1.2 Data used for Brix and pH prediction

The data for Brix and pH prediction were gathered from 85 mangoes. The obtained data comprised Lab and RGB color data of mangoes at various stages of the ripening process, softmax values, Brix values, and pH values for each mango. The softmax layer values were produced by classifying mango images with the VGG16 model trained on the data reported in Section 3.11. The Brix and pH values of the juice in each stage of mango ripening were measured from class 0 to class 5, which consisted of 12, 9, 12, 17, 19, and 12 mangoes, respectively. The lab and RGB data were obtained from three places on the top, middle, and bottom of a mango. Each mango's collected values included nine Lab values, nine RGB values, six softmax values, one Brix value, and one pH value.

3.2 Research Instruments

Python, OpenCV, Tensorflow, and Keras were used to create CNN, MobileNetV2, ResNet50, and VGG16 models. Mango images were taken using a PULUZ 30 x 30 x 30 cm lighting studio shooting tent box equipped with about 25 lumen LED brightness and a 5500 Kelvin color temperature of white light. The photos were taken with a OnePlus Ace2 camera. Color, Brix, and pH values were measured using a colorimeter, a refractometer, and a pH meter, respectively. Each mango was peeled and blended using a blending machine. Then, mash the blended mango and put it in glass tubes. After that, it was centrifuged to extract mango juice. Table 3.1 shows the instruments utilized in this study.

Table 3.1 Instruments

No.	Instruments	Purpose
1.	Photo box (PULUZ): Model 30 x 30 x 30 cm.	Box for taking mango photos
2.	Centrifuge (SURYQ): Model 800D	Contrifuge mashed mango to extract their juice
3.	Colorimeter (Linshang): Model LS171 and LScolor app installed from the Google play	Measure RGB and Lab color values
4.	Refractometer (ATAGO): Model PAL-1	Measure Brix values
5.	pH tester: Model BLE-C66	Measure pH and salinity values from juice
6.	OnePlus Ace2	Take mango photos

3.2.1 Mango sample preparation

To comprehensively study the maturity stages of mangoes, we bought them from Simummuang Market, there are two sets of mango samples at various maturity stages.

Figure 3.1 Mangoes from the market.

Source: Researcher, 2024

3.2.2 Image data collection

A photo box shown in Figure 3.2 was used for taking mango photos. This ensured consistent lighting and background conditions. For this data collection, we collected 1342 images, ranging from class 0 to class 5, with numbers 198, 156, 294, 350, 172, and 172, respectively.

Figure 3.2 Collecting an image of a mango.

Source: Researcher, 2024

3.2.3 Color data collection

For each mango, the colors were measured from three sampling points (the top, middle, and bottom sections of a mango) using a colorimeter, as shown in Figures 3.3-3.4

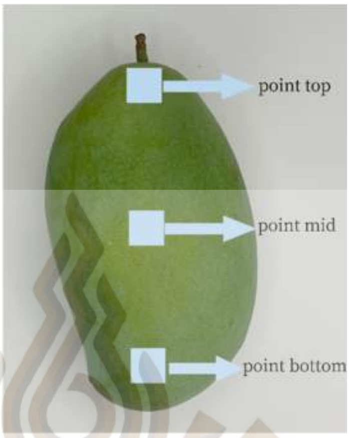

Figure 3.3 Color selection points.

Source: Researcher, 2024

Figure 3.4 Collect color data.

Source: Researcher, 2024

At each point, RGB and Lab color values were recorded. This process was repeated for all mangoes in the sample set. The data obtained were shown in Figure 3.5. Collect three points at the top, middle, and bottom of each mango, each containing three color data points. Each mango contains 9 RGB values and 9 Lab values.

R1	G1	B1	R2	G2	B2	R3	G3	B3
138	141	93	122	124	81	120	125	80
134	136	80	130	134	84	126	129	80
153	148	93	146	148	101	134	135	84

L1	a1	b1	L2	a2	b2	L3	a3	b3
57. 2	-8. 95	24. 89	50. 93	-8. 23	23. 14	50. 91	-9. 07	23. 54
55. 42	-9. 38	29. 54	54. 41	-9. 18	26. 03	52. 57	-8. 54	25. 44
60. 56	-6. 4	29. 26	60. 09	-8. 23	24. 26	55. 22	-8. 72	27. 05
58. 42	-10. 05	28. 18	54. 85	-9. 51	25. 93	56. 38	-8. 96	25. 83

Figure 3.5 Color data.

Source: Researcher, 2024

3.2.4 Brix and pH collection

A different set of mangoes at various maturity stages was prepared for the data analysis. Each mango was peeled, crushed, and placed in a centrifuge to extract the juice. After peeling the mangoes, the mango flesh was placed in a blender and blended until it formed a smooth pulp. Transfer the mango pulp into a test tube, using a funnel-shaped glass apparatus and a glass rod to ensure a clean transfer of all the pulp. Place the test tube in a centrifuge set to 3000 revolutions per minute (rpm) and let it spin for 20 minutes. After the centrifugation process is complete, carefully remove the test tube from the centrifuge. The clear mango juice will have separated from the solids. Figure 3.6 shows the process of extracting mango juice.

Figure 3.6 Obtaining fruit juice through a centrifuge.

Source: Researcher, 2024

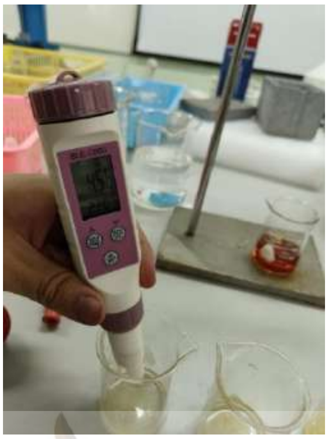

Figure 3.7 Measure pH.

Source: Researcher, 2024

Then a digital refractometer, as shown in Figure 3.6, was used to measure the Brix value, indicating the sugar content of the juice. Figure 3.7 shows the use of a pH meter to measure the pH value of the extracted juice.

Figure 3.8 Measure Brix.

Source: Researcher, 2024

All collected data were meticulously recorded and organized for subsequent analysis.

3.4 Data Analysis

3.4.1 Image classification using machine learning classsifiers

Figure 3.9 shows the image processing flow of the first part of the study. After collecting mango images, preprocessing is first performed(resize mango images), and then convolutional neural networks CNN, VGG16, ResNet, and MobilenNet are used to train and classify the images. At the same time, use the softmax layer data of the mango image output from VGG16 with good classification results for subsequent research.

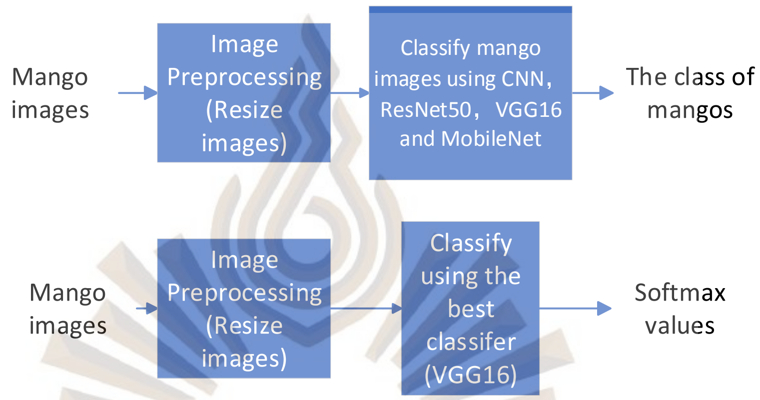

Figure 3.9 Image processing flow.

Source: Researcher, 2024

Here are the four neural networks which be used:

The CNNs comprised of 3 convolution layers. Each layer used 32 filters to extract image features. The flatten layer transformed the features into a one-dimensional vector passed through the ANN. The ANN, which was a part of the CNN, consisted of 64 hidden nodes and 6 output nodes.

ResNet-50 was utilized. The top section of the ResNet-50 featured a global average pooling layer, a dense layer with 64 hidden nodes, a 0.2 dropout layer, a dense layer with 64 hidden nodes and a ReLu activation function, and a 0.2 dropout layer.

MobileNetV3 was implemented. The top section of the MobileNetV3 featured a global average pooling layer, a dense layer with 64 hidden nodes, a 0.2 dropout layer, a dense layer with 64 hidden nodes and a ReLu activation function, and a 0.2 dropout layer.

VGG16 was applied. A global average pooling layer, a dense layer with 64 hidden nodes, a 0.2 dropout layer, a dense layer with 64 hidden nodes and a ReLu activation function, and a 0.2 dropout layer made up the top portion of the VGG16 that was employed.

The selection of these four network methods was driven by the need to explore and compare their performance across different architectures, ensuring robustness in various scenarios. The CNN serves as a foundational network capable of extracting basic features from images. ResNet-50, with its residual connections, addresses the vanishing gradient problem in deep networks, enhancing learning capability and stability. MobileNetV3, designed for efficiency, is suitable for resource-constrained environments while maintaining strong performance. VGG16, though a classic and simpler deep network, still excels in many image classification tasks. By employing these four diverse architectures, a comprehensive evaluation of the strengths and weaknesses of each method is possible, providing data to support the final model selection.

3.4.2 Predicting Brix and pH from softmax, RGB and Lab values

Figure 3.10 shows the data processing flow for the second part of the study. The color data of mango, RGB, Lab, and softmax layer data collected through a colorimeter and a convolutional neural network VGG16, were used as mango features. The relationship between these features and Brix and pH was explored using linear regression equations and ANN.

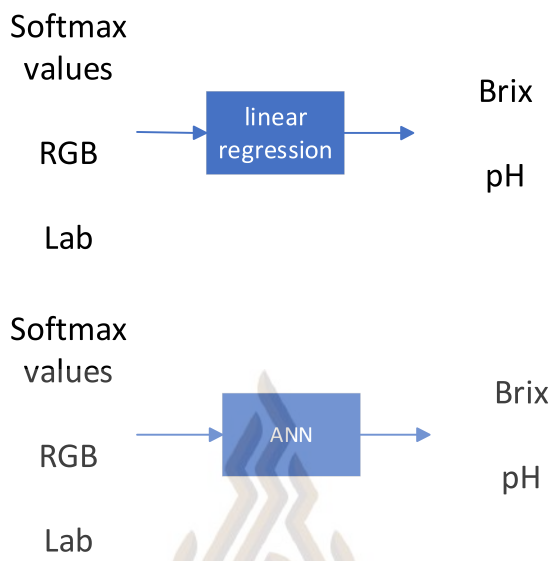

Figure 3.10 Predicting Brix and pH from softmax, RGB and Lab values.

Source: Researcher, 2024

The ANN architecture used in the analysis consists of an input layer with nodes corresponding to the number of features, two hidden layers each with 32 units, and an output layer with six nodes for classification. The model was trained to predict pH and Brix values, with Table 3.3 detailing the specific configuration. This includes an input layer, two hidden layers with 32 units each using ReLU or Sigmoid activation functions, and an output layer with a single unit for regression.

Table 3.2 ANN architecture

Layer	Type	Number of Neurons	Activation Function
Input Layer	Input	-	-
Hidden Layer 1	Dense	32	Sigmoid or ReLU
Output Layer	Dense	1	(None)

The predictive results from linear regression and ANN models were compared and analyzed to evaluate the efficacy of each method in predicting the chemical properties of mangoes based on their color and image data.

The R^2 value measures the goodness-of-fit between the actual and predicted values, ranging from 0 to 1. A higher R^2 value indicates a better fit, meaning the model explains a greater proportion of the variance in the actual data.

3.4.3 Result analysis tool

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{3-1}$$

To calcualate the classification accuracy of each class, a multiclass confusion matrix was converted into a binary confusion matrix for each class, and Equation 3.1 was used to calculate the accuracy. Equations 3.2, 3.3, and 3.4 demonstrate the method of calculating errors after predicting the Brix or pH values of mango juice using linear regression equations or ANN.

$$\text{error=Prediction values-Actual values} \tag{3-2}$$

$$\text{Absolute error=}|\text{error}| \tag{3-3}$$

$$\text{Relative error=}\frac{\text{Absolute error}}{\text{Actual value}} \tag{3-4}$$

$$\text{Percentage error=Relative error}*100\% \tag{3-5}$$

In the analysis of research results, the standard deviation (SD) and the coefficient of determination (R^2) are commonly used statistical metrics. The standard deviation measures the dispersion of data, indicating how much the data points deviate from the mean. The R^2 value, on the other hand, assesses the goodness of fit of a model, representing the proportion of variance in the dependent variable that is explained by the independent variables. An R^2 value closer to 1 indicates a better model fit and higher predictive accuracy. These metrics are crucial in evaluating the accuracy and reliability of the model in the analysis.

Chapter 4

Research Results

The results are divided into four sections: Section 4.1 focuses on image classification results using convolutional neural networks, with Subsection 4.1.1 utilizing an 80% training set and a 20% testing set, and Subsection 4.1.2 employing a 70% training set and a 30% testing set. Section 4.2 presents the prediction results between Brix and pH values using softmax data obtained from the VGG16 network. Sections 4.3 and 4.4 show the predicted results for Brix and pH values, with Section 4.3 using 9 RGB color values from 3 color points and Section 4.4 using 9 Lab color values from 3 color points, both obtained directly from a colorimeter.

4.1 Image classification results

This section contains the first part of the research findings, which focuses on the results obtained from training and predicting images of mangoes at different stages using different convolutional neural networks.

4.1.1 80% images for training and 20% images for testing

Table 4.1 shows the confusion matrix when using CNN. Many errors were found in the classification of class 4 and class 2 mangoes, indicating that CNN could not accurately distinguish the six different ripeness levels of mangoes

Table 4.1 Confusion matrix obtained by classifying the ripeness of mangoes using CNN (80% of images for training)

Result / Actual class	*Class 0*	*Class 1*	*Class 2*	*Class 3*	*Class 4*	*Class 5*
Class 0	86.84% (33)	13.16% (5)	0%	0%	0%	0%
Class 1	7.69% (2)	92.31% (24)	0%	0%	0%	0%
Class 2	0%	4.69% (3)	68.75% (44)	26.56% (17)	0%	0%
Class 3	0%	0%	0%	100% (70)	0%	0%
Class 4	0%	0%	0%	6.25% (2)	31.25% (10)	62.5% (20)
Class 5	0%	0%	0%	0%	0%	100% (32)

According to the results shown in Table 4.2, many errors were found when using MobileNet, revealing that MobileNet could not accurately classify mango ripeness levels. The percentage that correctly classified mangoes into classes 4 and 5 was only 78.13% and 75%, respectively.

Table 4.2 Confusion matrix obtained by classifying the ripeness of mangoes using MobileNet (80% of images for training)

Result / Actual class	*Class 0*	*Class 1*	*Class 2*	*Class 3*	*Class 4*	*Class 5*
Class 0	92.11% (35)	7.89% (3)	0%	0%	0%	0%
Class 1	11.54% (3)	88.46% (23)	0%	0%	0%	0%
Class 2	0%	4.69% (3)	92.19% (59)	3.13% (2)	0%	0%
Class 3	0%	0%	11.43% (8)	81.43% (57)	7.14% (5)	0%
Class 4	0%	0%	0%	0%	78.13% (25)	21.88% (7)
Class 5	0%	0%	0%	9.38% (3)	15.63% (5)	75% (24)

Table 4.3 shows the confusion matrix of classification with ResNet50. Significant errors were found in several classes, especially for the mangoes in classes 1 and 4. The percentage of correctly classifying class 1 mangoes was only 26.92%, and more than 50% of class 4 mangoes were misclassified as class 5 mangoes.

Table 4.3 Confusion matrix obtained by classifying the ripeness of mangoes using ResNet50 (80% of images for training)

Result / Actual class	*Class 0*	*Class 1*	*Class 2*	*Class 3*	*Class 4*	*Class 5*
Class 0	86.84% (33)	13.16% (5)	0%	0%	0%	0%
Class 1	34.62% (9)	26.92% (7)	38.46% (10)	0%	0%	0%
Class 2	1.56% (1)	4.69% (3)	92.19% (59)	1.56% (1)	0%	0%
Class 3	0%	0%	4.29% (3)	71.43% (50)	20% (14)	4.29% (3)
Class 4	0%	0%	0%	3.13% (1)	43.75% (14)	53.13% (17)
Class 5	0%	0%	0%	0%	6.25% (2)	93.75% (30)

Based on the confusion matrix in Table 4.4, for each ripeness class, an accuracy of more than 80% could be achieved, indicating that VGG16, trained on 80% of the images, was able to correctly predict the different ripeness levels of mangoes.

Table 4.4 Confusion matrix obtained by classifying the ripeness of mangoes using VGG16 (80% of images for training)

Result / Actual class	*Class 0*	*Class 1*	*Class 2*	*Class 3*	*Class 4*	*Class 5*
Class 0	92.11% (35)	7.89% (3)	0%	0%	0%	0%
Class 1	0%	100% (26)	0%	0%	0%	0%
Class 2	0%	0%	82.81% (53)	17.19% (11)	0%	0%
Class 3	0%	0%	0%	100% (70)	0%	0%
Class 4	0%	0%	0%	0%	81.25% (26)	18.75% (6)
Class 5	0%	0%	0%	0%	18.75% (6)	81.25% (26)

From the results obtained in Tables 1 to 4, VGG16 demonstrated high accuracy in classifying mango ripeness levels, which was superior to the other three networks, while ResNet50 performed poorly in classification.

4.1.2 70% images for training and 30% images for testing

Table 4.5 shows the confusion matrix of classification using CNN. Significant errors were found in several classes, particularly in class 4 mangoes. More than 70% of the class 4 mangoes were incorrectly classified as class 5 mangoes.

Table 4.5 Confusion matrix obtained by classifying the ripeness of mangoes using CNN (70% of images for training)

Result / Actual class	*Class 0*	*Class 1*	*Class 2*	*Class 3*	*Class 4*	*Class 5*
Class 0	70.69% (41)	29.31% (17)	0%	0%	0%	0%
Class 1	19.44% (7)	58.33% (21)	22.22% (8)	0%	0%	0%
Class 2	0%	10.58% (11)	83.65% (87)	5.77% (6)	0%	0%
Class 3	0%	0%	13.64% (15)	78.18% (86)	8.18% (9)	0%
Class 4	0%	0%	0%	13.46% (7)	15.38% (8)	71.15% (37)
Class 5	0%	0%	0%	0%	28.57% (12)	71.43% (30)

Table 4.6 presents the confusion matrix when using MobileNet. Major errors were found in classifying class 4 mangoes. More than 80% of the class 4 mangoes were incorrectly classified as class 5 mangoes.

Table 4.6 Confusion matrix obtained by classifying the ripeness of mangoes using MobileNet (70% of images for training)

Result / Actual class	*Class 0*	*Class 1*	*Class 2*	*Class 3*	*Class 4*	*Class 5*
Class 0	93.1% (54)	6.9% (4)	0%	0%	0%	0%
Class 1	11.11% (4)	88.89% (32)	0%	0%	0%	0%
Class 2	0%	3.85% (4)	96.15% (100)	0%	0%	0%
Class 3	0%	0%	9.09% (10)	90.91% (100)	0%	0%
Class 4	0%	0%	0%	0%	17.31% (9)	82.69% (43)
Class 5	0%	0%	0%	0%	0%	100% (42)

Table 4.7 depicts the confusion matrix when using ResNet50. Several significant errors were found in several classes, preventing ResNet50 from attaining high accuracy.

Table 4.7 Confusion matrix obtained by classifying the ripeness of mangoes using ResNet50 (70% of images for training)

Actual class \ Result	*Class 0*	*Class 1*	*Class 2*	*Class 3*	*Class 4*	*Class 5*
Class 0	86.21% (50)	13.79% (8)	0%	0%	0%	0%
Class 1	41.67% (15)	33.33% (12)	25% (9)	0%	0%	0%
Class 2	0%	8.65% (9)	50.96% (53)	40.38% (42)	0%	0%
Class 3	0%	0%	1.82% (2)	71.82% (79)	24.55% (27)	1.82% (2)
Class 4	0%	0%	0%	38.46% (20)	5.77% (3)	55.77% (29)
Class 5	0%	0%	0%	4.76% (2)	2.38% (1)	92.86% (39)

Table 4.8 illustrates the confusion matrix when using VGG16. Since there were insufficient images to train VGG16, the classification accuracy of class 4 mangoes was low. In class 4, 44.13% of mangoes were correctly classified. However, Table 4.4 shows a significant increase in the percentage to 81.25% when we used 80% of the images for training.

Table 4.8 Confusion matrix obtained by classifying the ripeness of mangoes using VGG16 (70% of images for training)

Actual class \ Result	*Class 0*	*Class 1*	*Class 2*	*Class 3*	*Class 4*	*Class 5*
Class 0	96.55% (56)	3.45% (2)	0%	0%	0%	0%
Class 1	5.56% (2)	94.44% (34)	0%	0%	0%	0%
Class 2	0%	3.85% (4)	96.15% (100)	0%	0%	0%
Class 3	0%	0%	0.91% (1)	96.36% (106)	2.73% (3)	0%
Class 4	0%	0%	0%	0%	44.23% (23)	55.77% (29)
Class 5	0%	0%	0%	0%	4.76% (2)	95.24% (40)

The results showed that VGG16 performed better than CNN, MobileNet and ResNet50. Using 80% of the images for training achieved higher accuracies than training with 70% of the images. Apart from the types of classifiers, the number of training data had a significant impact on the accuracy of mango ripeness classification.

Table 4.9 summarizes and compares the results when using different machine learning classifiers and images for training.

Table 4.9 Classification accuracy

Classifier \ Accuracy	Accuracy (80% of images for training)	Accuracy (70% of images for training)
CNN	81.30%	67.91%
MobileNet	85.11%	83.83%
ResNet50	73.66%	58.71%
VGG16	90.08%	89.30%

4.2 Brix and pH prediction results using softmax values

This part of the chapter explored the use of classification results to determine Brix and pH values. The softmax layer values from the classification were used to predict Brix and pH values using linear regression and ANN. The study was divided into six parts, each focusing on the relationship between different data types and Brix or pH values, using both prediction methods. It investigated the relationships between softmax values, RGB color data, and Lab color values with Brix and pH levels. Each part provided insights into the effectiveness of linear regression and ANN for predicting Brix and pH values.

4.2.1 Brix prediction using linear regression

The outcomes, standard deviation, and error of the predictions against the actual values when predicting Brix values from softmax values are displayed in Tables 4.10–4.12. The outcomes of the Brix value prediction process using softmax values are shown in Table 4.10. The analysis yielded the following linear regression equation.

$$Brix = 12.322 - soft\,max\,1 * 4.735 - soft\,max\,2 * 2.921 - soft\,max\,3 * 2.278 + softmax5 * 2.08 + softmax6 * 3.799 \quad (4.1)$$

Using the softmax and Brix values for linear regression, the R^2 was 0.959, suggesting that the equation could accurately predict the outcome.

Table 4.10 Predicting Brix values from softmax values using linear regression

Mango No.	Class	Actual Brix Value	Prediction Brix Value	Error	Absolute error	Relative error	Percentage error
1	class0	7.4	7.59	0.19	0.19	0.03	2.55
2	class0	8	7.59	-0.41	0.41	0.05	5.14
3	class0	7.3	7.60	0.30	0.30	0.04	4.10
4	class0	8.1	7.60	-0.50	0.50	0.06	6.22
5	class0	7.8	7.59	-0.21	0.21	0.03	2.72
6	class0	7.1	7.59	0.49	0.49	0.07	6.87
7	class0	7.9	7.59	-0.31	0.31	0.04	3.94
8	class0	7.1	7.59	0.49	0.49	0.07	6.89
9	class0	6.8	7.59	0.79	0.79	0.12	11.61
10	class0	7.7	7.59	-0.11	0.11	0.01	1.45
11	class0	7.8	7.59	-0.21	0.21	0.03	2.72
12	class0	8.2	7.68	-0.52	0.52	0.06	6.37
13	class1	9.8	9.40	-0.40	0.40	0.04	4.08
14	class1	8.9	9.40	0.50	0.50	0.06	5.62
15	class1	9.3	9.40	0.10	0.10	0.01	1.06
16	class1	9.4	9.40	0.00	0.00	0.00	0.01
17	class1	9.3	9.41	0.11	0.11	0.01	1.13
18	class1	9.8	9.40	-0.40	0.40	0.04	4.10
19	class1	9.7	9.40	-0.30	0.30	0.03	3.04
20	class1	9.3	9.40	0.10	0.10	0.01	1.11
21	class1	9	9.41	0.41	0.41	0.05	4.57
22	class2	11	9.99	-1.01	1.01	0.09	9.23
23	class2	9.6	10.03	0.43	0.43	0.05	4.51
24	class2	11.2	10.04	-1.16	1.16	0.10	10.32
25	class2	9.7	10.04	0.34	0.34	0.04	3.54
26	class2	9.2	10.04	0.84	0.84	0.09	9.17
27	class2	9.6	10.04	0.44	0.44	0.05	4.62
28	class2	10.3	10.04	-0.26	0.26	0.02	2.49
29	class2	9.2	10.04	0.84	0.84	0.09	9.18
30	class2	10.2	10.05	-0.15	0.15	0.02	1.51
31	class2	11.4	10.04	-1.36	1.36	0.12	11.89
32	class2	9.7	10.04	0.34	0.34	0.04	3.54

Table 4.10 Predicting Brix values from softmax values using linear regression(cont)

Mango No.	Class	Actual Brix Value	Prediction Brix Value	Error	Absolute error	Relative error	Percentage error
33	class2	9.5	10.04	0.54	0.54	0.06	5.73
34	class2	10	10.04	0.04	0.04	0.00	0.44
35	class3	11.8	12.32	0.52	0.52	0.04	4.42
36	class3	12.2	12.32	0.12	0.12	0.01	1.00
37	class3	12.9	12.32	-0.58	0.58	0.04	4.48
38	class3	12.2	12.32	0.12	0.12	0.01	1.00
39	class3	11.9	12.32	0.42	0.42	0.04	3.55
40	class3	11.9	12.32	0.42	0.42	0.04	3.54
41	class3	12.9	12.32	-0.58	0.58	0.04	4.49
42	class3	11.7	12.32	0.62	0.62	0.05	5.32
43	class3	12	12.32	0.32	0.32	0.03	2.68
44	class3	12.2	12.32	0.12	0.12	0.01	1.00
45	class3	13	12.32	-0.68	0.68	0.05	5.21
46	class3	11.8	12.32	0.52	0.52	0.04	4.42
47	class3	12.7	12.32	-0.38	0.38	0.03	2.98
48	class3	12.9	12.32	-0.58	0.58	0.04	4.48
49	class3	11.8	12.32	0.52	0.52	0.04	4.42
50	class3	12.5	12.32	-0.18	0.18	0.01	1.42
51	class3	12.8	12.32	-0.48	0.48	0.04	3.73
52	class3	12.6	12.32	-0.28	0.28	0.02	2.20
53	class4	15.3	14.40	-0.90	0.90	0.06	5.89
54	class4	14.3	14.40	0.10	0.10	0.01	0.70
55	class4	13.3	14.40	1.10	1.10	0.08	8.27
56	class4	15.4	14.65	-0.75	0.75	0.05	4.90
57	class4	15	14.64	-0.36	0.36	0.02	2.38
58	class4	13.8	14.55	0.75	0.75	0.05	5.40
59	class4	14.4	14.62	0.22	0.22	0.02	1.56
60	class4	14.3	14.43	0.13	0.13	0.01	0.90
61	class4	14.5	14.43	-0.07	0.07	0.01	0.50
62	class4	13.4	14.42	1.02	1.02	0.08	7.65
63	class4	15	14.41	-0.59	0.59	0.04	3.91
64	class4	14.7	14.42	-0.28	0.28	0.02	1.88
65	class4	14.3	14.42	0.12	0.12	0.01	0.87
66	class4	14.9	14.44	-0.46	0.46	0.03	3.10
67	class4	14.7	14.44	-0.26	0.26	0.02	1.76
68	class4	14.5	14.45	-0.05	0.05	0.00	0.36
69	class4	13.6	14.40	0.80	0.80	0.06	5.89
70	class4	15.5	14.40	-1.10	1.10	0.07	7.08
71	class4	13.3	14.40	1.10	1.10	0.08	8.29
72	class4	14	15.81	1.81	1.81	0.13	12.93

Table 4.10 Predicting Brix values from softmax values using linear regression(cont)

Mango No.	Class	Actual Brix Value	Prediction Brix Value	Error	Absolute error	Relative error	Percentage error
73	class5	16.7	15.55	-1.15	1.15	0.07	6.88
74	class5	15.7	15.58	-0.12	0.12	0.01	0.79
75	class5	16.6	15.38	-1.22	1.22	0.07	7.33
76	class5	16.5	16.12	-0.38	0.38	0.02	2.33
77	class5	15.8	16.12	0.32	0.32	0.02	2.00
78	class5	16.2	16.12	-0.08	0.08	0.01	0.52
79	class5	16	16.11	0.11	0.11	0.01	0.71
80	class5	15.6	16.11	0.51	0.51	0.03	3.29
81	class5	16.3	16.11	-0.19	0.19	0.01	1.15
82	class5	16.2	16.11	-0.09	0.09	0.01	0.53
83	class5	16.7	16.11	-0.59	0.59	0.04	3.53
84	class5	15.6	16.11	0.51	0.51	0.03	3.27
85	class5	16.1	16.11	0.01	0.01	0.00	0.07

Table 4.11 displays the absolute Brix prediction error for each class of mangoes, relative error, percentage error, and corresponding standard deviation. Table 4.12 shows the average actual and predicted values of mango classes, as well as their standard deviations. In Table 4.11, class 1 had an absolute error of 0.2574, a relative error of 0.0275, and a percentage error of 2.7476%. In contrast, class 2 exhibited higher errors, with an absolute error of 0.5982, a relative error of 0.0586, and a percentage error of 5.8594%. Class 1 demonstrates lower and more evenly distributed errors compared to class 2, indicating more stable predictions. Overall, the average errors across all classes are 0.4628 for absolute error, 0.0398 for relative error, and 3.9817% for percentage error.

Table 4.11 Absolute and relative error of Brix predictions (linear regression and softmax values)

Class	class0	class1	class2	class3	class4	class5	All classes
Absolute error	0.3781	0.2574	0.5982	0.4136	0.5990	0.4065	0.4628
Standard deviation	0.1916	0.1814	0.4067	0.1836	0.4750	0.3927	0.3539
Relative error	0.0505	0.0275	0.0586	0.0335	0.0421	0.0249	0.0398
Standard deviation	0.0278	0.0197	0.0369	0.0148	0.0343	0.0236	0.0292
Percentage error	5.0466	2.7476	5.8594	3.3534	4.2113	2.4920	3.9817
Standard deviation	2.7806	1.9654	3.6904	1.4763	3.4250	2.3591	2.9201

Table 4.12 compares the average actual and predicted Brix values across different classes. Class1 and class2 exhibited relatively similar average actual Brix values (9.3889 vs. 10.0462), suggesting comparable characteristics within these classes. Notably, class 3 showed a similar match between predicted and actual Brix values, indicating highly accurate prediction in this class. However, we found variations in standard deviations across classes, especially in class 4 and class 5, which indicated greater variability in predicted Brix values compared to others.

Table 4.12 Average and standard deviation of actual and predicted Brix values (linear regression and softmax values)

	class0	class1	class2	class3	class4	class5
Average Brix (Actual values)	7.6000	9.3889	10.0462	12.3222	14.4100	16.1538
Standard deviation	0.4492	0.3257	0.7378	0.4634	0.6805	0.3992
Average Brix (Predicted values)	7.5975	9.4026	10.0386	12.3220	14.4597	15.9607
Standard deviation	0.0256	0.0042	0.0164	0.0003	0.0856	0.2638

Figure 4.1 shows the prediction of Brix values from softmax values using linear regression. The graphs show the actual and predicted Brix values when using linear regression.

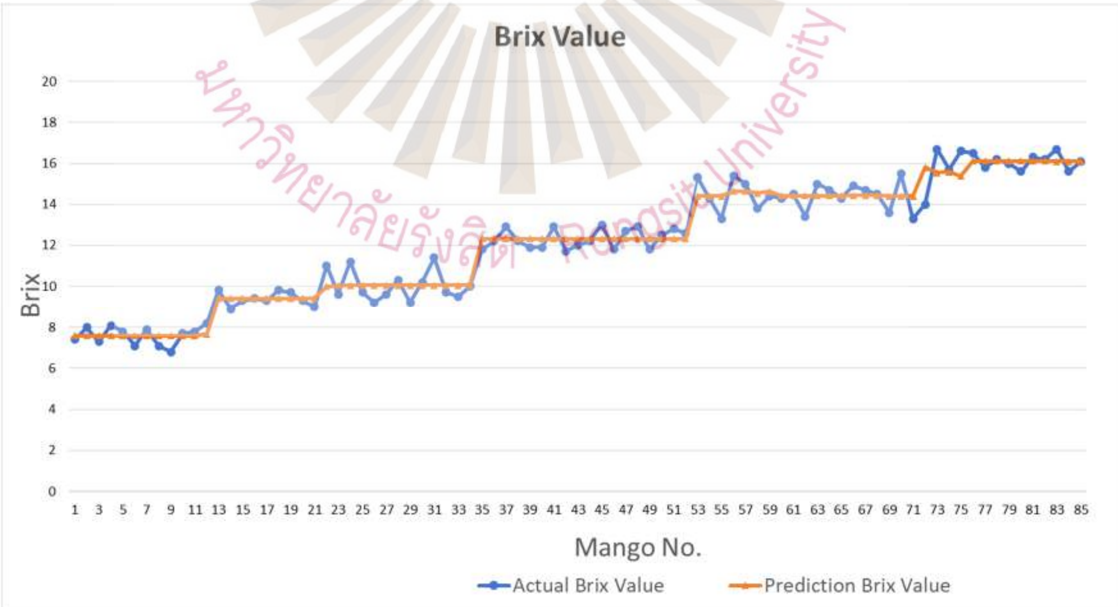

Figure 4.1 Graphs shown the actual and predicted Brix values when predicting using linear regression softmax value.

Source: Researcher, 2024

4.2.2 Brix prediction using ANN

Table 4.13 displays the results of predicting Brix values based on softmax values when using an Artificial Neural Network (ANN) for regression. The table highlights the model's performance, showcasing a high degree of accuracy and low error metrics, which underscore the effectiveness of the ANN in this predictive task.

Table 4.13 Results of predicting Brix values from softmax values using ANN

Mango No.	Class	Actual Brix Value	Prediction Brix Value	Error	Absolute error	Relative error	Percentage error
1	class0	8	7.98	-0.02	0.02	0.00	0.24
2	class0	7.3	7.87	0.57	0.57	0.08	7.87
3	class0	8.1	8.16	0.06	0.06	0.01	0.72
4	class0	7.8	8.90	1.10	1.10	0.14	14.12
5	class0	7.1	8.25	1.15	1.15	0.16	16.19
6	class0	7.9	7.83	-0.07	0.07	0.01	0.90
7	class0	7.1	7.30	0.20	0.20	0.03	2.79
8	class0	6.8	7.49	0.69	0.69	0.10	10.15
9	class0	7.7	7.38	-0.32	0.32	0.04	4.21
10	class0	7.8	7.69	-0.11	0.11	0.01	1.38
11	class0	8.2	7.67	-0.53	0.53	0.06	6.42
12	class0	9.8	8.71	-1.09	1.09	0.11	11.12
13	class1	8.9	9.47	0.57	0.57	0.06	6.44
14	class1	9.3	9.15	-0.15	0.15	0.02	1.58
15	class1	9.4	9.14	-0.26	0.26	0.03	2.77
16	class1	9.3	9.30	0.00	0.00	0.00	0.01
17	class1	9.8	8.66	-1.14	1.14	0.12	11.61
18	class1	9.7	9.26	-0.44	0.44	0.05	4.59
19	class1	9.3	8.95	-0.35	0.35	0.04	3.77
20	class1	9	9.33	0.33	0.33	0.04	3.63
21	class1	11	10.07	-0.93	0.93	0.08	8.47
22	class2	9.6	9.79	0.19	0.19	0.02	1.95
23	class2	11.2	10.27	-0.93	0.93	0.08	8.32
24	class2	9.7	10.22	0.52	0.52	0.05	5.40
25	class2	9.2	9.69	0.49	0.49	0.05	5.36
26	class2	9.6	10.39	0.79	0.79	0.08	8.23
27	class2	10.3	10.41	0.11	0.11	0.01	1.07
28	class2	9.2	10.03	0.83	0.83	0.09	9.04
29	class2	10.2	9.84	-0.36	0.36	0.04	3.56
30	class2	11.4	10.32	-1.08	1.08	0.10	9.51
31	class2	9.7	9.71	0.01	0.01	0.00	0.09
32	class2	9.5	9.73	0.23	0.23	0.02	2.39

Table 4.13 Results of predicting Brix values from softmax values using ANN(cont)

Mango No.	Class	Actual Brix Value	Prediction Brix Value	Error	Absolute error	Relative error	Percentage error
33	class2	10	9.75	-0.25	0.25	0.02	2.46
34	class2	11.8	12.22	0.42	0.42	0.04	3.60
35	class3	12.2	12.18	-0.02	0.02	0.00	0.20
36	class3	12.9	12.26	-0.64	0.64	0.05	5.00
37	class3	12.2	12.05	-0.15	0.15	0.01	1.22
38	class3	11.9	12.37	0.47	0.47	0.04	3.96
39	class3	11.9	12.32	0.42	0.42	0.04	3.56
40	class3	12.9	12.28	-0.62	0.62	0.05	4.79
41	class3	11.7	12.20	0.50	0.50	0.04	4.23
42	class3	12	12.50	0.50	0.50	0.04	4.14
43	class3	12.2	12.47	0.27	0.27	0.02	2.24
44	class3	13	12.20	-0.80	0.80	0.06	6.12
45	class3	11.8	12.23	0.43	0.43	0.04	3.65
47	class3	12.9	12.41	-0.49	0.49	0.04	3.76
48	class3	11.8	12.29	0.49	0.49	0.04	4.14
49	class3	12.5	12.51	0.01	0.01	0.00	0.05
50	class3	12.8	12.35	-0.45	0.45	0.04	3.51
51	class3	12.6	12.07	-0.53	0.53	0.04	4.19
52	class3	15.3	14.41	-0.89	0.89	0.06	5.82
53	class4	14.3	15.03	0.73	0.73	0.05	5.12
54	class4	13.3	14.16	0.86	0.86	0.06	6.49
55	class4	15.4	14.02	-1.38	1.38	0.09	8.99
56	class4	15	13.93	-1.07	1.07	0.07	7.16
57	class4	13.8	14.54	0.74	0.74	0.05	5.38
58	class4	14.4	14.25	-0.15	0.15	0.01	1.02
59	class4	14.3	14.52	0.22	0.22	0.02	1.51
60	class4	14.5	14.93	0.43	0.43	0.03	2.99
61	class4	13.4	14.72	1.32	1.32	0.10	9.87
62	class4	15	14.32	-0.68	0.68	0.05	4.55
63	class4	14.7	14.42	-0.28	0.28	0.02	1.90
64	class4	14.3	14.82	0.52	0.52	0.04	3.65
65	class4	14.9	14.19	-0.71	0.71	0.05	4.79
66	class4	14.7	14.70	0.00	0.00	0.00	0.00
67	class4	14.5	13.98	-0.52	0.52	0.04	3.56
68	class4	13.6	14.07	0.47	0.47	0.03	3.49
69	class4	15.5	14.96	-0.54	0.54	0.03	3.49
70	class4	13.3	14.59	1.29	1.29	0.10	9.67
71	class4	14	13.77	-0.23	0.23	0.02	1.61
72	class4	16.7	16.24	-0.46	0.46	0.03	2.73
73	class5	15.7	16.32	0.62	0.62	0.04	3.96

Table 4.13 Results of predicting Brix values from softmax values using ANN(cont)

Mango No.	Class	Actual Brix Value	Prediction Brix Value	Error	Absolute error	Relative error	Percentage error
74	class5	16.6	16.21	-0.39	0.39	0.02	2.34
75	class5	16.5	16.21	-0.29	0.29	0.02	1.78
76	class5	15.8	16.18	0.38	0.38	0.02	2.38
77	class5	16.2	16.42	0.22	0.22	0.01	1.34
78	class5	16	16.06	0.06	0.06	0.00	0.40
79	class5	15.6	16.16	0.56	0.56	0.04	3.57
80	class5	16.3	16.47	0.17	0.17	0.01	1.04
81	class5	16.2	16.01	-0.19	0.19	0.01	1.17
82	class5	16.7	16.05	-0.65	0.65	0.04	3.91
83	class5	15.6	15.93	0.33	0.33	0.02	2.15
84	class5	16.1	16.18	0.08	0.08	0.00	0.48
85	class5	16.1	16.54	0.44	0.44	0.03	2.73

Table 4.14 displays the absolute Brix prediction error for each class of mangoes, relative error, percentage error, and corresponding standard deviation, when using ANN for the regression. Table 4.14 provides error metrics for various mango classes. Class 4 exhibited the highest absolute error at 0.6307, whereas class 5 showed the lowest at 0.3368. The average absolute error across all classes was 0.4841. Standard deviations for absolute error vary significantly, with class 0 and class 4 having the highest at 0.4327 and 0.3958, respectively. In terms of relative error, class 5 demonstrates the lowest at 0.0209, whereas class 0 has the highest at 0.0634. The average relative error across all classes was 0.0420. Regarding percentage error, class 5 showed the lowest variability at 2.0946%, while class 0 showed the highest at 6.3415%. On average, the percentage error is 4.2022% across all classes. These metrics illustrated the varying levels of accuracy and variability in error among different classes.

Table 4.15 shows the average actual and predicted Brix values of mango classes, as well as their standard deviations. The average actual Brix values ranged from 7.8000 for class 0 to 16.1077 for class 5, demonstrating a progressive increase across classes. The variability in actual Brix values is notable, with class 2 having the highest standard deviation at 0.8490 and class 5 having the lowest at 0.3639. Comparatively, predicted Brix values generally aligned closely with actual values across all classes, with minor deviations. However, we found variability in the predicted Brix values, especially in class 2, which had the highest standard deviation of 0.6736, and class 5, which had the lowest at 0.1755. This analysis underscored both the accuracy of predicted Brix values and the variability within each class, providing insights into prediction reliability across different Brix levels.

Table 4.14 Absolute and relative error of Brix predictions (ANN and softmax values)

Class	class0	class1	class2	class3	class4	class5	Average
Absolute error	0.4925	0.4637	0.4786	0.4363	0.6307	0.3368	0.4841
Standard deviation	0.4327	0.3663	0.3381	0.2380	0.3958	0.1941	0.3390
Relative error	0.0634	0.0476	0.0469	0.0345	0.0440	0.0209	0.0420
Standard deviation	0.0554	0.0358	0.0321	0.0175	0.0284	0.0121	0.0330
Percentage error	6.3415	4.7635	4.6909	3.4455	4.3995	2.0946	4.2022
Standard deviation	5.5388	3.5784	3.2061	1.7511	2.8398	1.2119	3.3021

Table 4.15 Average and standard deviation of actual and predicted Brix values (ANN and softmax values)

	class0	class1	class2	class3	class4	class5
Average Brix(Actual values)	7.8000	9.5222	10.1077	12.5167	14.4800	16.1077
Standard deviation	0.7711	0.6241	0.8490	0.8248	0.8320	0.3639
Average Brix (Predicted values)	7.9362	9.2584	10.1824	12.4231	14.4173	16.2127
Standard deviation	0.4994	0.3846	0.6736	0.5152	0.3804	0.1755

Figure 4.2 shows the prediction of Brix values from softmax values using ANN for linear regression. The overall prediction results showed that ANN for regression provided better prediction results than linear regression.

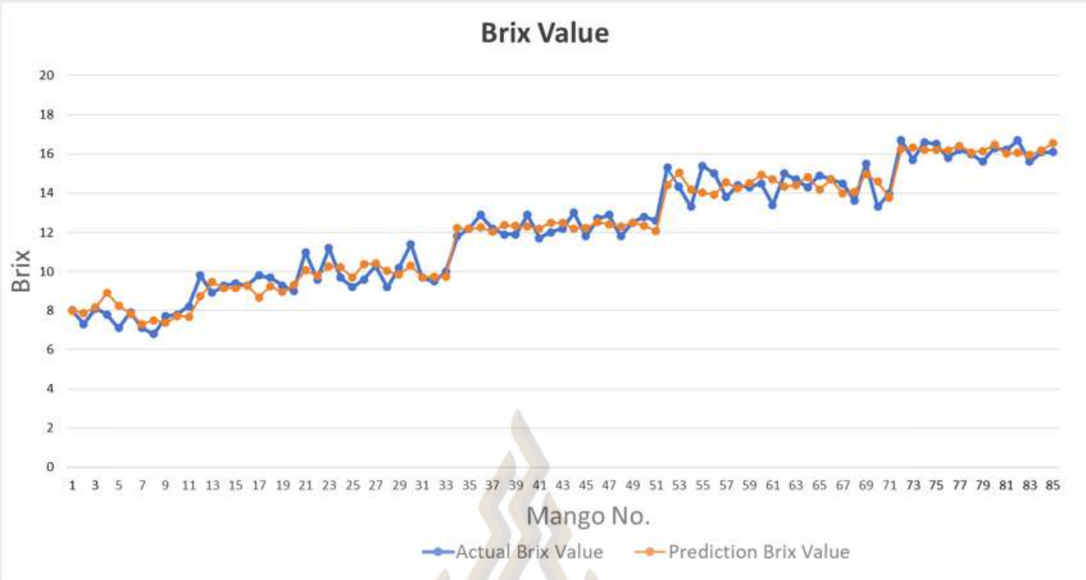

Figure 4.2 Graphs shown the actual and predicted Brix values when predicting using ANN and softmax values.
Source: Researcher, 2024

Tables 4.10, 4.11, 4.12, 4.13, 4.14, and 4.15 show the relationship between the softmax values obtained from VGG16 classifiers and mango juice sweetness, which are predicted by linear regression equations and ANN. Tables 4.11 and 4.14 show the error values between predicted values and actual values. The result showed that ANN for regression could perform better than linear regression in predicting Brix values from the softmax values.

4.2.3 pH prediction using linear regression

Tables 4.16–4.18 show the results, error, and standard deviation of predictions versus actual values when predicting pH values from softmax values. Table 4.16 displays the results of predicting Brix values based on softmax values. After applying linear regression, the obtained linear formula was as follows:

$$pH = 2.868 - soft\,max\,1 * 0.855 - soft\,max\,2 * 0.623 - soft\,max\,3 * 0.218 + softmax5 * 0.347 + softmax6 * 1.145 \quad (4\text{-}2)$$

The R^2 obtained through linear regression using the softmax and pH values was 0.915, indicating the equation could be used for the prediction.

Table 4.16 Results of predicting pH values from softmax values using linear regression

Mango No.	Class	Actual pH Value	Prediction pH Value	Error	Absolute error	Relative error	Percentage error
1	class0	2.08	2.01	-0.07	0.07	0.03	3.37
2	class0	1.97	2.01	0.04	0.04	0.02	2.03
3	class0	2.02	2.01	-0.01	0.01	0.00	0.50
4	class0	2.08	2.01	-0.07	0.07	0.03	3.37
5	class0	1.97	2.01	0.04	0.04	0.02	2.03
6	class0	1.96	2.01	0.05	0.05	0.03	2.55
7	class0	2.09	2.01	-0.08	0.08	0.04	3.83
8	class0	1.91	2.01	0.10	0.1	0.05	5.24
9	class0	1.99	2.01	0.02	0.02	0.01	1.01
10	class0	1.99	2.01	0.02	0.02	0.01	1.01
11	class0	2.04	2.01	-0.03	0.03	0.01	1.47
12	class0	2.08	2.02	-0.06	0.06	0.03	2.88
13	class1	2.11	2.25	0.14	0.14	0.07	6.64
14	class1	2.3	2.25	-0.05	0.05	0.02	2.17
15	class1	2.58	2.25	-0.33	0.33	0.13	12.79
16	class1	2.07	2.25	0.18	0.18	0.09	8.70
17	class1	2.17	2.25	0.08	0.08	0.04	3.69
18	class1	2.24	2.24	0.00	0	0.00	0.00
19	class1	2.01	2.25	0.24	0.24	0.12	11.94
20	class1	2.59	2.25	-0.34	0.34	0.13	13.13
21	class1	2.14	2.25	0.11	0.11	0.05	5.14
22	class2	2.7	2.61	-0.09	0.09	0.03	3.33
23	class2	2.48	2.64	0.16	0.16	0.06	6.45
24	class2	2.82	2.65	-0.17	0.17	0.06	6.03
25	class2	2.74	2.65	-0.09	0.09	0.03	3.28
26	class2	2.9	2.65	-0.25	0.25	0.09	8.62
27	class2	2.63	2.65	0.02	0.02	0.01	0.76
28	class2	2.64	2.65	0.01	0.01	0.00	0.38
29	class2	2.89	2.65	-0.24	0.24	0.08	8.30
30	class2	2.9	2.65	-0.25	0.25	0.09	8.62
31	class2	2.32	2.65	0.33	0.33	0.14	14.22
32	class2	2.67	2.65	-0.02	0.02	0.01	0.75
33	class2	2.38	2.65	0.27	0.27	0.11	11.34
34	class2	2.34	2.65	0.31	0.31	0.13	13.25
35	class3	3.05	2.87	-0.18	0.18	0.06	5.90
36	class3	2.79	2.87	0.08	0.08	0.03	2.87
37	class3	3.07	2.87	-0.20	0.2	0.07	6.51
38	class3	2.96	2.87	-0.09	0.09	0.03	3.04
39	class3	2.91	2.87	-0.04	0.04	0.01	1.37
40	class3	2.76	2.87	0.11	0.11	0.04	3.99

Table 4.16 Results of predicting pH values from softmax values using linear regression(cont)

Mango No.	Class	Actual pH Value	Prediction pH Value	Error	Absolute error	Relative error	Percentage error
41	class3	2.93	2.87	-0.06	0.06	0.02	2.05
42	class3	2.72	2.87	0.15	0.15	0.06	5.51
43	class3	2.83	2.87	0.04	0.04	0.01	1.41
44	class3	2.73	2.87	0.14	0.14	0.05	5.13
45	class3	2.75	2.87	0.12	0.12	0.04	4.36
46	class3	2.91	2.87	-0.04	0.04	0.01	1.37
47	class3	2.96	2.87	-0.09	0.09	0.03	3.04
48	class3	2.95	2.87	-0.08	0.08	0.03	2.71
49	class3	2.87	2.87	0.00	0	0.00	0.00
50	class3	2.95	2.87	-0.08	0.08	0.03	2.71
51	class3	2.79	2.87	0.08	0.08	0.03	2.87
52	class3	2.7	2.87	0.17	0.17	0.06	6.30
53	class4	2.95	3.21	0.26	0.26	0.09	8.81
54	class4	3.35	3.22	-0.13	0.13	0.04	3.88
55	class4	3.25	3.21	-0.04	0.04	0.01	1.23
56	class4	3.17	3.33	0.16	0.16	0.05	5.05
57	class4	3.39	3.33	-0.06	0.06	0.02	1.77
58	class4	3.37	3.28	-0.09	0.09	0.03	2.67
59	class4	3.44	3.32	-0.12	0.12	0.03	3.49
60	class4	2.88	3.23	0.35	0.35	0.12	12.15
61	class4	3.42	3.23	-0.19	0.19	0.06	5.56
62	class4	3.42	3.23	-0.19	0.19	0.06	5.56
63	class4	3.31	3.22	-0.09	0.09	0.03	2.72
64	class4	2.95	3.23	0.28	0.28	0.09	9.49
65	class4	3.43	3.23	-0.20	0.2	0.06	5.83
66	class4	3.28	3.23	-0.05	0.05	0.02	1.52
67	class4	3.13	3.23	0.10	0.1	0.03	3.19
68	class4	3.4	3.24	-0.16	0.16	0.05	4.71
69	class4	3.05	3.22	0.17	0.17	0.06	5.57
70	class4	2.85	3.22	0.37	0.37	0.13	12.98
71	class4	3.37	3.22	-0.15	0.15	0.04	4.45
72	class4	3.15	3.87	0.72	0.72	0.23	22.86
73	class5	4.14	3.75	-0.39	0.39	0.09	9.42
74	class5	4.23	3.76	-0.47	0.47	0.11	11.11
75	class5	3.83	3.67	-0.16	0.16	0.04	4.18
76	class5	4.15	4.01	-0.14	0.14	0.03	3.37
77	class5	4.25	4.01	-0.24	0.24	0.06	5.65
78	class5	4.13	4.01	-0.12	0.12	0.03	2.91

Table 4.16 Results of predicting pH values from softmax values using linear regression(cont)

Mango No.	Class	Actual pH Value	Prediction pH Value	Error	Absolute error	Relative error	Percentage error
79	class5	3.96	4.01	0.05	0.05	0.01	1.26
80	class5	3.82	4.01	0.19	0.19	0.05	4.97
81	class5	3.83	4.01	0.18	0.18	0.05	4.70
82	class5	3.88	4.01	0.13	0.13	0.03	3.35
83	class5	4.09	4.01	-0.08	0.08	0.02	1.96
84	class5	3.95	4.01	0.06	0.06	0.02	1.52
85	class5	3.94	4.01	0.07	0.07	0.02	1.78

Table 4.17 presents the absolute pH prediction error for each class of mangoes, relative error, and percentage error. Class 0 demonstrated the smallest absolute error at 0.0492, while class 1 showed the highest at 0.1633. The average absolute error across all classes was 0.1433. The variation in absolute error is notable, with class 4 having the highest standard deviation at 0.1541 and class 0 having the lowest at 0.0275. In terms of relative error, class 0 boasts the lowest at 0.0244, contrasting with class 1's highest at 0.0713. The average relative error stood at 0.0494. Regarding percentage error, class 0 exhibited the lowest variability at 2.4389%, whereas class 1 showed the highest at 7.1322%. The average percentage error across all classes was 4.9368%, with class 4 having the highest standard deviation in percentage error (5.1147%) and class 0 having the lowest (1.3764%).

Table 4.17 Absolute and relative error of pH predictions (linear regression and softmax values)

Class	class0	class1	class2	class3	class4	class5	Average
Absolute error	0.0492	0.1633	0.1700	0.0972	0.1940	0.1754	0.1433
Standard deviation	0.0275	0.1199	0.1142	0.0542	0.1541	0.1269	0.1193
Relative error	0.0244	0.0713	0.0657	0.0340	0.0617	0.0432	0.0494
Standard deviation	0.0138	0.0481	0.0470	0.0188	0.0511	0.0299	0.0403
Percentage error	2.4389	7.1322	6.5653	3.3975	6.1748	4.3210	4.9368
Standard deviation	1.3764	4.8097	4.6987	1.8848	5.1147	2.9940	4.0338

Table 4.18 shows the average actual and predicted pH values of mango classes, as well as their standard deviations. The average actual pH values ranged from 2.0150

for class 0 to 4.0154 for class 5, showing a progressive increase. Variability in actual pH values was notable, with class 1 having the highest standard deviation at 0.2107 and class 3 having the lowest at 0.1138. The standard deviation for predicted pH values indicated variation across classes, with class 4 showing the highest at 0.0398 and class 3 the lowest at 0.0000314, highlighting the prediction model's accuracy and variability across different classes.

Table 4.18 Average values and standard deviation of actual and predicted pH values (linear regression and softmax values)

	class0	class1	class2	class3	class4	class5
Average pH (Actual values)	2.0150	2.2456	2.6469	2.8683	3.2280	4.0154
Standard deviation	0.0590	0.2107	0.2101	0.1138	0.1986	0.1559
Average pH (Predicted values)	2.0108	2.2489	2.6462	2.8700	3.2437	3.9393
Standard deviation	0.0029	0.0033	0.0112	0.0000 314	0.0398	0.1226

Figure 4.3 shows the prediction of pH values from softmax values using linear regression.

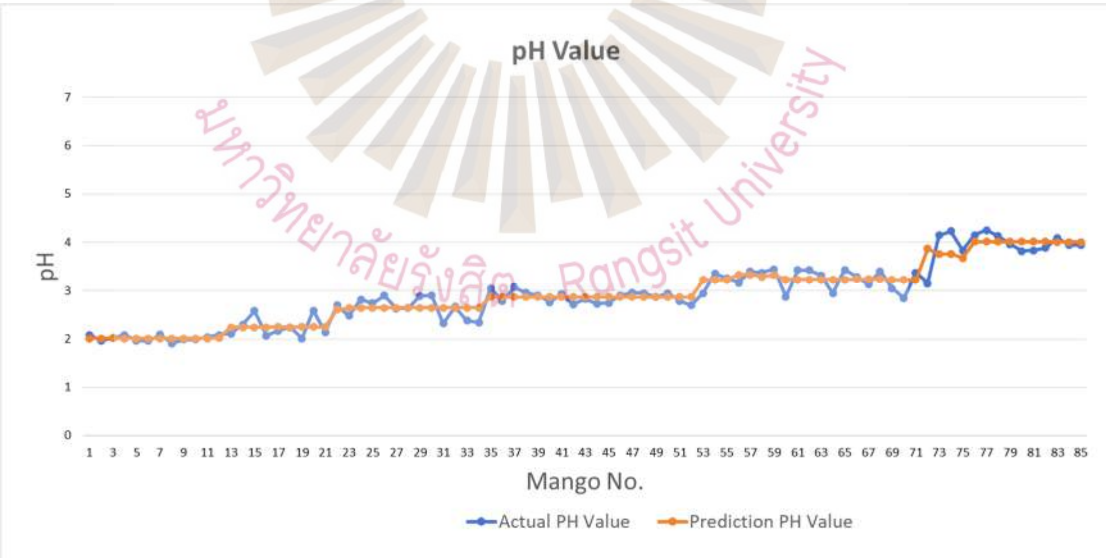

Figure 4.3 Graphs shown the actual and predicted pH values when predicting using linear regression softmax values.

Source: Researcher, 2024

4.2.4 pH prediction results using ANN

Table 4.19 presents the absolute pH prediction error for each class of mangoes, relative error, percentage error.

Table 4.19 Results of predicting pH values from softmax values using ANN

Mango No.	Class	Actual pH Value	Prediction pH Value	Error	Absolute error	Relative error	Percentage error
1	class0	2.08	2.08	0.00	0.00	0.00	0.03
2	class0	1.97	2.07	0.10	0.10	0.05	5.33
3	class0	2.02	2.02	0.00	0.00	0.00	0.15
4	class0	2.08	2.09	0.01	0.01	0.01	0.59
5	class0	1.97	2.11	0.14	0.14	0.07	7.05
6	class0	1.96	1.95	-0.01	0.01	0.00	0.35
7	class0	2.09	1.96	-0.13	0.13	0.06	6.41
8	class0	1.91	1.97	0.06	0.06	0.03	3.23
9	class0	1.99	1.96	-0.03	0.03	0.01	1.32
10	class0	1.99	1.93	-0.06	0.06	0.03	2.93
11	class0	2.04	1.96	-0.08	0.08	0.04	4.09
12	class0	2.08	2.15	0.07	0.07	0.03	3.48
13	class1	2.11	2.27	0.16	0.16	0.07	7.36
14	class1	2.3	2.31	0.01	0.01	0.01	0.65
15	class1	2.58	2.25	-0.33	0.33	0.13	12.61
16	class1	2.07	2.26	0.19	0.19	0.09	9.18
17	class1	2.17	2.18	0.01	0.01	0.01	0.64
19	class1	2.01	2.23	0.22	0.22	0.11	10.83
20	class1	2.59	2.33	-0.26	0.26	0.10	10.05
21	class1	2.14	2.69	0.55	0.55	0.26	25.85
22	class2	2.7	2.57	-0.13	0.13	0.05	4.72
23	class2	2.48	2.62	0.14	0.14	0.06	5.79
24	class2	2.82	2.69	-0.13	0.13	0.05	4.53
25	class2	2.74	2.59	-0.15	0.15	0.06	5.56
26	class2	2.9	2.71	-0.19	0.19	0.07	6.57
27	class2	2.63	2.78	0.15	0.15	0.06	5.70
28	class2	2.64	2.68	0.04	0.04	0.02	1.58
29	class2	2.89	2.57	-0.32	0.32	0.11	11.17
30	class2	2.9	2.68	-0.22	0.22	0.08	7.56
31	class2	2.32	2.61	0.29	0.29	0.12	12.44
32	class2	2.67	2.65	-0.02	0.02	0.01	0.81
33	class2	2.38	2.60	0.22	0.22	0.09	9.24
34	class2	2.34	2.91	0.57	0.57	0.24	24.41
35	class3	3.05	2.91	-0.14	0.14	0.05	4.68
36	class3	2.79	2.96	0.17	0.17	0.06	6.21
37	class3	3.07	2.78	-0.29	0.29	0.10	9.54

Table 4.19 Results of predicting pH values from softmax values using ANN(cont)

Mango No.	Class	Actual pH Value	Prediction pH Value	Error	Absolute error	Relative error	Percentage error
38	class3	2.96	2.93	-0.03	0.03	0.01	1.05
39	class3	2.91	2.88	-0.03	0.03	0.01	1.07
40	class3	2.76	2.89	0.13	0.13	0.05	4.88
41	class3	2.93	2.79	-0.14	0.14	0.05	4.70
42	class3	2.72	2.91	0.19	0.19	0.07	6.86
43	class3	2.83	2.94	0.11	0.11	0.04	3.95
44	class3	2.73	2.79	0.06	0.06	0.02	2.21
45	class3	2.75	2.82	0.07	0.07	0.02	2.37
46	class3	2.91	2.91	0.00	0.00	0.00	0.07
47	class3	2.96	2.87	-0.09	0.09	0.03	3.08
48	class3	2.95	2.82	-0.13	0.13	0.04	4.46
49	class3	2.87	2.84	-0.03	0.03	0.01	1.13
50	class3	2.95	2.80	-0.15	0.15	0.05	4.97
51	class3	2.79	2.85	0.06	0.06	0.02	2.28
52	class3	2.7	3.17	0.47	0.47	0.18	17.57
53	class4	2.95	3.27	0.32	0.32	0.11	10.97
54	class4	3.35	3.15	-0.20	0.20	0.06	5.99
55	class4	3.25	3.30	0.05	0.05	0.01	1.41
56	class4	3.17	3.29	0.12	0.12	0.04	3.87
57	class4	3.39	3.31	-0.08	0.08	0.02	2.38
58	class4	3.37	3.32	-0.05	0.05	0.02	1.51
59	class4	3.44	3.21	-0.23	0.23	0.07	6.69
60	class4	2.88	3.29	0.41	0.41	0.14	14.15
61	class4	3.42	3.24	-0.18	0.18	0.05	5.15
62	class4	3.42	3.19	-0.23	0.23	0.07	6.79
63	class4	3.31	3.22	-0.09	0.09	0.03	2.76
64	class4	2.95	3.26	0.31	0.31	0.11	10.56
65	class4	3.43	3.20	-0.23	0.23	0.07	6.81
66	class4	3.28	3.26	-0.02	0.02	0.01	0.62
67	class4	3.13	3.16	0.03	0.03	0.01	1.09
68	class4	3.4	3.13	-0.27	0.27	0.08	7.87
69	class4	3.05	3.27	0.22	0.22	0.07	7.34
70	class4	2.85	3.22	0.37	0.37	0.13	13.12
71	class4	3.37	3.64	0.27	0.27	0.08	7.95
72	class4	3.15	3.89	0.74	0.74	0.24	23.53
73	class5	4.14	3.92	-0.22	0.22	0.05	5.28
74	class5	4.23	3.84	-0.39	0.39	0.09	9.30
75	class5	3.83	4.03	0.20	0.20	0.05	5.18
76	class5	4.15	3.99	-0.16	0.16	0.04	3.85
77	class5	4.25	4.03	-0.22	0.22	0.05	5.07

Table 4.19 Results of predicting pH values from softmax values using ANN(cont)

Mango No.	Class	Actual pH Value	Prediction pH Value	Error	Absolute error	Relative error	Percentage error
78	class5	4.13	3.99	-0.14	0.14	0.03	3.41
79	class5	3.96	4.03	0.07	0.07	0.02	1.68
80	class5	3.82	4.03	0.21	0.21	0.06	5.58
81	class5	3.83	3.93	0.10	0.10	0.03	2.51
82	class5	3.88	4.00	0.12	0.12	0.03	3.20
83	class5	4.09	3.93	-0.16	0.16	0.04	3.96
84	class5	3.95	3.98	0.03	0.03	0.01	0.79
85	class5	3.94	3.98	0.04	0.04	0.01	1.02

Table 4.20 shows Absolute errors for different classes ranged from 0.2574 (class 1) to 0.5990 (class 4), with an overall average of 0.3119. The highest variability in absolute error was observed in class 2, with a standard deviation of 0.4067, while class 0 showed the lowest variability at 0.0499. Relative errors ranged from 0.0275 (class 1) to 0.0540 (class 2), with an average of 0.0398. The highest variability in relative error was noted in class 1, with a standard deviation of 0.0796, and the lowest in class 5, with a standard deviation of 0.0228. Percentage errors varied between 2.7476% (class 1) and 5.8594% (class 2), averaging 3.9817%. Class 1 had the highest standard deviation in percentage errors at 7.9560, while class 5 had the lowest at 2.2814. These metrics indicate significant differences in pH prediction accuracy across different classes, suggesting that certain classes, particularly class 2, may benefit from further refinement in the model.

Table 4.20 Absolute and relative error of pH predictions (ANN and softmax values)

Class	class0	class1	class2	class3	class4	class5	Average
Absolute error	0.3781	0.2574	0.5982	0.4136	0.5990	0.5800	0.3119
Standard deviation	0.0499	0.1814	0.4067	0.1836	0.4750	0.4750	0.1387
Relative error	0.0505	0.0275	0.0540	0.0335	0.0421	0.0249	0.0398
Standard deviation	0.0248	0.0796	0.0601	0.0404	0.0554	0.0228	0.0515
Percentage error	5.0466	2.7476	5.8594	3.3534	4.2113	2.4920	3.9817
Standard deviation	2.4756	7.9560	6.0099	4.0354	5.5388	2.2814	5.1842

Table 4.21 compares the average actual and predicted pH values across different classes. The actual pH values ranged from 2.0150 for class 0 to 4.0154 for class 5, showing a progressive increase across classes. Class 3 had the lowest standard

deviation in actual pH values at 0.1138, indicating less variability compared to other classes. The predicted pH values closely aligned with actual values, with minor deviations noted across all classes. Class 0 exhibited the highest standard deviation in predicted pH values at 0.0758, suggesting slightly greater variability in predictions compared to other classes. Overall, the prediction model could effectively capture the trend in pH values across different classes.

Table 4.21 Average values and standard deviation of actual and predicted pH values (ANN and softmax values)

	class0	class1	class2	class3	class4	class5
Average pH (Actual values)	2.0150	2.2456	2.6469	2.8683	3.2280	4.0154
Standard deviation	0.0590	0.2107	0.2101	0.1138	0.1986	0.1559
Average pH (Predicted values)	2.0221	2.3051	2.6664	2.8814	3.2599	3.9693
Standard deviation	0.0758	0.1525	0.0960	0.0924	0.1068	0.0601

Figure 4.4 shows the comparison between ANN-predicted pH values and actual values. The graphs showed that in classes 4 and 5 of mangoes, the prediction performance was not very good.

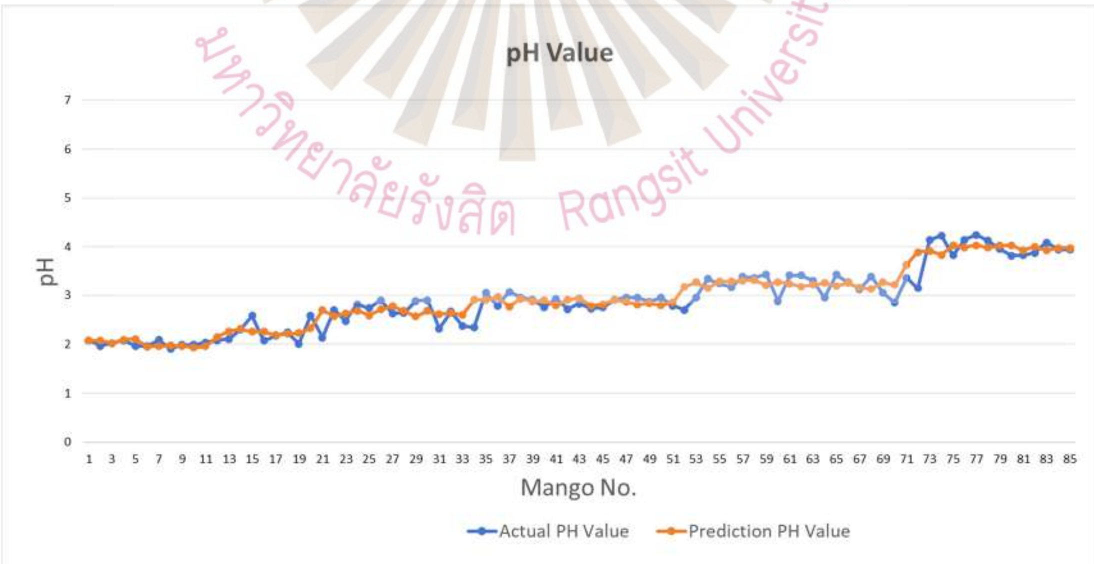

Figure 4.4 Graphs shown the actual and predicted pH values when predicting using ANN softmax values.

Source: Researcher, 2024

Tables 4.16, 4.17, 4.18, 4.19, 4.20, and 4.21 show how the values from a VGG16 softmax layer are related to the pH value of mango juice, as predicted by the linear regression equation and ANN. Tables 4.17 and 4.20 show the average error values between predicted and actual values. The linear regression had a lower average error than the ANN (0.1433 vs. 0.3119).

4.3 Brix and pH prediction results using RGB

4.3.1 Brix prediction results using linear regression

After applying linear regression to predict Brix values, the obtained equation was as follows:

$$Brix = -7.591 + R1 * 0.032 - G1 * 0.033 + B1 * 0.021 + R2 * 0.023 + G2 * 0.008 + B2 * 0.008 + R3 * 0.002 + G3 * 0.066 - B3 * 1.571E - 005 \quad (4.3)$$

The linear regression of RGB and Brix had an R^2 of 0.935, which could provide accurate predictions.

Table 4.22 displays the errors in predicting Brix values based on RGB values measured from three points.

Table 4.22 Predicting Brix values from RGB values using linear regression

Mango No.	Class	Actual Brix Value	Prediction Brix Value	Error	Absolute error	Relative error	Percentage error
1	class0	8	7.06	-0.94	0.94	0.12	11.75
2	class0	7.3	7.39	0.09	0.09	0.01	1.23
3	class0	8.1	8.90	0.80	0.8	0.10	9.88
4	class0	7.8	7.67	-0.13	0.13	0.02	1.67
5	class0	7.1	8.23	1.13	1.13	0.16	15.92
6	class0	7.9	8.09	0.19	0.19	0.02	2.41
7	class0	7.1	8.03	0.93	0.93	0.13	13.10
8	class0	6.8	7.81	1.01	1.01	0.15	14.85
9	class0	7.7	8.22	0.52	0.52	0.07	6.75
10	class0	7.8	8.11	0.31	0.31	0.04	3.97
11	class0	8.2	7.92	-0.28	0.28	0.03	3.41
12	class0	9.8	8.95	-0.85	0.85	0.09	8.67
13	class1	8.9	9.10	0.20	0.2	0.02	2.25
14	class1	9.3	8.59	-0.71	0.71	0.08	7.63

Table 4.22 Predicting Brix values from RGB values using linear regression(cont)

Mango No.	Class	Actual Brix Value	Prediction Brix Value	Error	Absolute error	Relative error	Percentage error
15	class1	9.4	7.84	-1.56	1.56	0.17	16.60
16	class1	9.3	7.97	-1.33	1.33	0.14	14.30
17	class1	9.8	10.18	0.38	0.38	0.04	3.88
18	class1	9.7	10.52	0.82	0.82	0.08	8.45
19	class1	9.3	10.48	1.18	1.18	0.13	12.69
20	class1	9	9.82	0.82	0.82	0.09	9.11
21	class1	11	9.06	-1.94	1.94	0.18	17.64
22	class2	9.6	10.59	0.99	0.99	0.10	10.31
23	class2	11.2	10.63	-0.57	0.57	0.05	5.09
24	class2	9.7	10.71	1.01	1.01	0.10	10.41
25	class2	9.2	10.33	1.13	1.13	0.12	12.28
26	class2	9.6	11.01	1.41	1.41	0.15	14.69
27	class2	10.3	10.65	0.35	0.35	0.03	3.40
28	class2	9.2	11.60	2.40	2.4	0.26	26.09
29	class2	10.2	8.97	-1.23	1.23	0.12	12.06
30	class2	11.4	10.35	-1.05	1.05	0.09	9.21
31	class2	9.7	10.07	0.37	0.37	0.04	3.81
32	class2	9.5	9.51	0.01	0.01	0.00	0.11
33	class2	10	9.65	-0.35	0.35	0.04	3.50
34	class2	11.8	9.79	-2.01	2.01	0.17	17.03
35	class3	12.2	13.86	1.66	1.66	0.14	13.61
36	class3	12.9	12.74	-0.16	0.16	0.01	1.24
37	class3	12.2	12.12	-0.08	0.08	0.01	0.66
38	class3	11.9	13.41	1.51	1.51	0.13	12.69
39	class3	11.9	11.87	-0.03	0.03	0.00	0.25
40	class3	12.9	11.93	-0.97	0.97	0.08	7.52
41	class3	11.7	12.80	1.10	1.1	0.09	9.40
42	class3	12	12.37	0.37	0.37	0.03	3.08
43	class3	12.2	12.21	0.01	0.01	0.00	0.08
44	class3	13	13.08	0.08	0.08	0.01	0.62
45	class3	11.8	13.52	1.72	1.72	0.15	14.58
46	class3	12.7	12.24	-0.46	0.46	0.04	3.62
47	class3	12.9	12.34	-0.56	0.56	0.04	4.34
48	class3	11.8	12.90	1.10	1.1	0.09	9.32
49	class3	12.5	12.33	-0.17	0.17	0.01	1.36
50	class3	12.8	13.12	0.32	0.32	0.02	2.50
51	class3	12.6	12.24	-0.36	0.36	0.03	2.86
52	class3	15.3	12.35	-2.95	2.95	0.19	19.28
53	class4	14.3	14.42	0.12	0.12	0.01	0.84
54	class4	13.3	14.55	1.25	1.25	0.09	9.40

Table 4.22 Predicting Brix values from RGB values using linear regression(cont)

Mango No.	Class	Actual Brix Value	Prediction Brix Value	Error	Absolute error	Relative error	Percentage error
55	class4	15.4	13.44	-1.96	1.96	0.13	12.73
56	class4	15	14.80	-0.20	0.2	0.01	1.33
57	class4	13.8	14.64	0.84	0.84	0.06	6.09
58	class4	14.4	14.59	0.19	0.19	0.01	1.32
59	class4	14.3	14.40	0.10	0.1	0.01	0.70
60	class4	14.5	14.68	0.18	0.18	0.01	1.24
61	class4	13.4	14.88	1.48	1.48	0.11	11.04
62	class4	15	14.43	-0.57	0.57	0.04	3.80
63	class4	14.7	14.69	-0.01	0.01	0.00	0.07
64	class4	14.3	13.97	-0.33	0.33	0.02	2.31
65	class4	14.9	14.67	-0.23	0.23	0.02	1.54
66	class4	14.7	14.39	-0.31	0.31	0.02	2.11
67	class4	14.5	14.27	-0.23	0.23	0.02	1.59
68	class4	13.6	13.80	0.20	0.2	0.01	1.47
69	class4	15.5	14.73	-0.77	0.77	0.05	4.97
70	class4	13.3	14.52	1.22	1.22	0.09	9.17
71	class4	14	14.71	0.71	0.71	0.05	5.07
72	class4	16.7	14.26	-2.44	2.44	0.15	14.61
73	class5	15.7	16.64	0.94	0.94	0.06	5.99
74	class5	16.6	16.92	0.32	0.32	0.02	1.93
75	class5	16.5	15.86	-0.64	0.64	0.04	3.88
76	class5	15.8	15.89	0.09	0.09	0.01	0.57
77	class5	16.2	16.47	0.27	0.27	0.02	1.67
78	class5	16	16.03	0.03	0.03	0.00	0.19
79	class5	15.6	16.29	0.69	0.69	0.04	4.42
80	class5	16.3	16.11	-0.19	0.19	0.01	1.17
81	class5	16.2	15.99	-0.21	0.21	0.01	1.30
82	class5	16.7	16.13	-0.57	0.57	0.03	3.41
83	class5	15.6	15.96	0.36	0.36	0.02	2.31
84	class5	16.1	16.65	0.55	0.55	0.03	3.42
85	class5	16.1	16.61	0.51	0.51	0.03	3.17

Table 4.23 indicates that Brix predictions generally align well with actual values, with average absolute errors ranging from 0.4131 to 0.9933. Class 1 shows the highest variability in absolute errors, while class 5 has the lowest. Relative errors vary from 0.0257 to 0.1028, with class 5 having the lowest variability. Percentage errors range from 2.5701% to 10.2828%, with class 5 again showing the lowest variability.

Overall, the model performs reasonably well but could be improved, particularly for classes with higher variability.

Table 4.23 Absolute and relative error of Brix predictions (using linear regression and RGB values)

Class	class0	class1	class2	class3	class4	class5	Average
Absolute error	0.5983	0.9933	0.9908	0.7561	0.6670	0.4131	0.7214
Standard deviation	0.3835	0.5603	0.6847	0.7944	0.6786	0.2634	0.6284
Relative error	0.0780	0.1028	0.0985	0.0594	0.0457	0.0257	0.0642
Standard deviation	0.0530	0.0540	0.0699	0.0587	0.0446	0.0166	0.0567
Percentage error	7.8012	10.2828	9.8456	5.9447	4.5699	2.5701	6.4231
Standard deviation	5.3015	5.3990	6.9904	5.8724	4.4558	1.6629	5.6698

Table 4.24 compares the actual and predicted average Brix values across different classes. Actual average Brix values ranged from 7.8000 (class 0) to 16.1077 (class 5), showing a progressive increase across classes. Standard deviations in actual Brix values ranged from 0.6241 (class 1) to 0.8490 (class 2). The predicted average Brix values were close to their actual values, with minor deviations observed across all classes. Class 1 exhibited the highest standard deviation in predicted Brix values at 1.0263, suggesting slightly greater variability in predictions compared to other classes.

Table 4.24 Average values and standard deviation of actual and predicted Brix values (using linear regression and RGB values)

	class0	class1	class2	class3	class4	class5
Average Brix(Actual values)	7.8000	9.5222	10.1077	12.5167	14.4800	16.1077
Standard deviation	0.7711	0.6241	0.8490	0.8248	0.8320	0.3639
Average Brix (Predicted values)	8.0317	9.2844	10.2969	12.6350	14.4516	16.1293
Standard deviation	0.5419	1.0263	0.6958	0.5738	0.3648	0.6332

Figure 4.4 shows the comparison between the Brix values predicted using linear regression and the actual values.

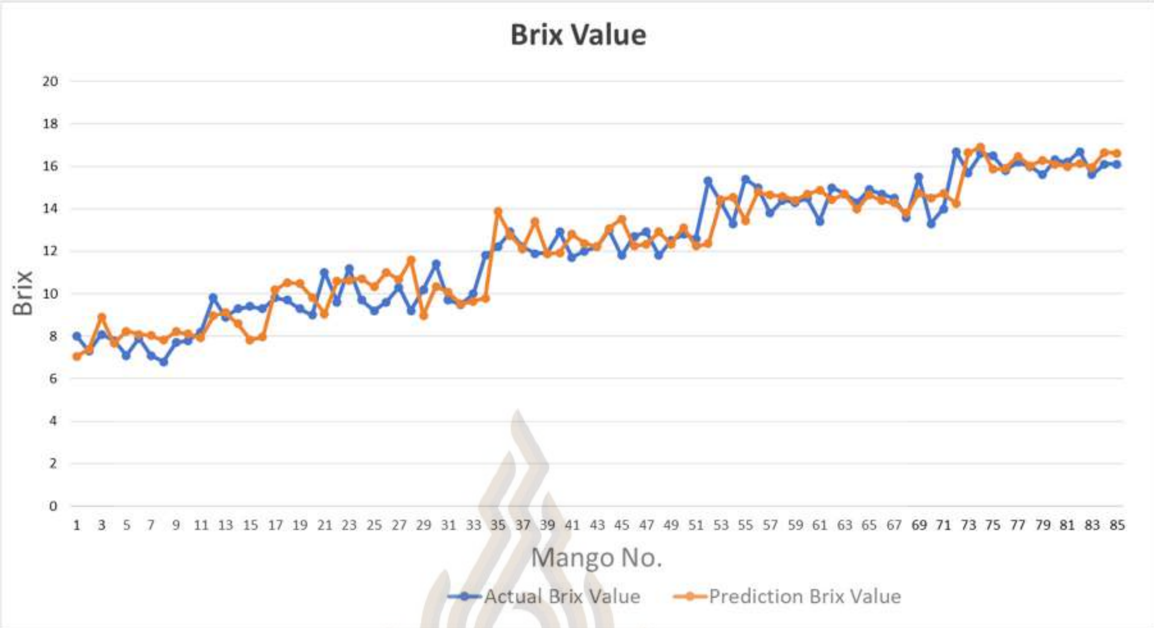

Figure 4.5 Graphs shown the actual and predicted Brix values when predicting using linear regression and RGB values.
Source: Researcher, 2024

4.3.2 Brix prediction results using ANN

Table 4.25 reveals the errors in predicting Brix values based on RGB values measured from three points using ANN.

Table 4.25 Predicting Brix values from softmax values using ANN

Mango No.	Class	Actual Brix Value	Prediction Brix Value	Error	Absolute error	Relative error	Percentage error
1	class0	8	7.58	-0.42	0.42	0.05	5.25
2	class0	7.3	8.48	1.18	1.18	0.16	16.16
3	class0	8.1	7.45	-0.65	0.65	0.08	8.02
4	class0	7.8	7.94	0.14	0.14	0.02	1.79
5	class0	7.1	8.05	0.95	0.95	0.13	13.38
6	class0	7.9	7.72	-0.18	0.18	0.02	2.28
7	class0	7.1	7.67	0.57	0.57	0.08	8.03
8	class0	6.8	7.73	0.93	0.93	0.14	13.68
9	class0	7.7	8.29	0.59	0.59	0.08	7.66
10	class0	7.8	7.82	0.02	0.02	0.00	0.26
11	class0	8.2	8.75	0.55	0.55	0.07	6.71
12	class0	9.8	8.99	-0.81	0.81	0.08	8.27
13	class1	8.9	9.03	0.13	0.13	0.01	1.46
14	class1	9.3	8.14	-1.16	1.16	0.12	12.47

Table 4.25 Predicting Brix values from softmax values using ANN(cont)

Mango No.	Class	Actual Brix Value	Prediction Brix Value	Error	Absolute error	Relative error	Percentage error
15	class1	9.4	8.07	-1.33	1.33	0.14	14.15
16	class1	9.3	9.79	0.49	0.49	0.05	5.27
17	class1	9.8	10.14	0.34	0.34	0.03	3.47
18	class1	9.7	10.17	0.47	0.47	0.05	4.85
19	class1	9.3	9.48	0.18	0.18	0.02	1.94
20	class1	9	8.50	-0.50	0.5	0.06	5.56
21	class1	11	10.01	-0.99	0.99	0.09	9.00
22	class2	9.6	10.46	0.86	0.86	0.09	8.96
23	class2	11.2	10.50	-0.70	0.7	0.06	6.25
24	class2	9.7	9.86	0.16	0.16	0.02	1.65
25	class2	9.2	10.73	1.53	1.53	0.17	16.63
26	class2	9.6	10.39	0.79	0.79	0.08	8.23
27	class2	10.3	11.50	1.20	1.2	0.12	11.65
28	class2	9.2	8.66	-0.54	0.54	0.06	5.87
29	class2	10.2	10.23	0.03	0.03	0.00	0.29
30	class2	11.4	9.81	-1.59	1.59	0.14	13.95
31	class2	9.7	9.27	-0.43	0.43	0.04	4.43
32	class2	9.5	9.31	-0.19	0.19	0.02	2.00
33	class2	10	9.54	-0.46	0.46	0.05	4.60
34	class2	11.8	13.67	1.87	1.87	0.16	15.85
35	class3	12.2	12.66	0.46	0.46	0.04	3.77
36	class3	12.9	12.23	-0.67	0.67	0.05	5.19
37	class3	12.2	13.47	1.27	1.27	0.10	10.41
38	class3	11.9	11.85	-0.05	0.05	0.00	0.42
39	class3	11.9	12.15	0.25	0.25	0.02	2.10
40	class3	12.9	12.46	-0.44	0.44	0.03	3.41
41	class3	11.7	12.03	0.33	0.33	0.03	2.82
42	class3	12	12.13	0.13	0.13	0.01	1.08
43	class3	12.2	12.38	0.18	0.18	0.01	1.48
44	class3	13	13.25	0.25	0.25	0.02	1.92
45	class3	11.8	12.12	0.32	0.32	0.03	2.71
46	class3	12.7	12.03	-0.67	0.67	0.05	5.28
47	class3	12.9	12.87	-0.03	0.03	0.00	0.23
48	class3	11.8	11.91	0.11	0.11	0.01	0.93
49	class3	12.5	12.80	0.30	0.3	0.02	2.40
50	class3	12.8	11.97	-0.83	0.83	0.06	6.48
51	class3	12.6	12.31	-0.29	0.29	0.02	2.30
52	class3	15.3	14.40	-0.90	0.9	0.06	5.88
53	class4	14.3	14.61	0.31	0.31	0.02	2.17
54	class4	13.3	13.49	0.19	0.19	0.01	1.43

Table 4.25 Predicting Brix values from softmax values using ANN(cont)

Mango No.	Class	Actual Brix Value	Prediction Brix Value	Error	Absolute error	Relative error	Percentage error
55	class4	15.4	14.52	-0.88	0.88	0.06	5.71
56	class4	15	14.71	-0.29	0.29	0.02	1.93
57	class4	13.8	14.53	0.73	0.73	0.05	5.29
58	class4	14.4	14.44	0.04	0.04	0.00	0.28
59	class4	14.3	14.69	0.39	0.39	0.03	2.73
60	class4	14.5	14.87	0.37	0.37	0.03	2.55
61	class4	13.4	14.53	1.13	1.13	0.08	8.43
62	class4	15	14.78	-0.22	0.22	0.01	1.47
63	class4	14.7	14.24	-0.46	0.46	0.03	3.13
64	class4	14.3	14.55	0.25	0.25	0.02	1.75
65	class4	14.9	14.31	-0.59	0.59	0.04	3.96
66	class4	14.7	14.34	-0.36	0.36	0.02	2.45
67	class4	14.5	13.96	-0.54	0.54	0.04	3.72
68	class4	13.6	14.65	1.05	1.05	0.08	7.72
69	class4	15.5	14.49	-1.01	1.01	0.07	6.52
70	class4	13.3	14.65	1.35	1.35	0.10	10.15
71	class4	14	14.31	0.31	0.31	0.02	2.21
72	class4	16.7	16.08	-0.62	0.62	0.04	3.71
73	class5	15.7	16.25	0.55	0.55	0.04	3.50
74	class5	16.6	15.53	-1.07	1.07	0.06	6.45
75	class5	16.5	15.47	-1.03	1.03	0.06	6.24
76	class5	15.8	16.12	0.32	0.32	0.02	2.03
77	class5	16.2	15.64	-0.56	0.56	0.03	3.46
78	class5	16	15.92	-0.08	0.08	0.01	0.50
79	class5	15.6	15.70	0.10	0.1	0.01	0.64
80	class5	16.3	15.50	-0.80	0.8	0.05	4.91
81	class5	16.2	15.74	-0.46	0.46	0.03	2.84
82	class5	16.7	15.60	-1.10	1.1	0.07	6.59
83	class5	15.6	16.14	0.54	0.54	0.03	3.46
84	class5	16.1	15.69	-0.41	0.41	0.03	2.55
85	class5	16.1	15.81	-0.29	0.29	0.02	1.80

Table 4.26 provides error metrics for pH prediction across different classes. Absolute errors ranged from 0.4156 (class 3) to 0.7962 (class 2), averaging 0.5742 across all classes. Class 2 showed the highest variability in absolute error with a standard deviation of 0.5895, while class 3 showed the lowest at 0.3327. Relative errors ranged from 0.0327 (class 3) to 0.0772 (class 2), with an average of 0.0507. Class 2 exhibited the highest variability in relative error, with a standard deviation of 0.0542, and class 5

had the lowest variability at 0.0207. Percentage errors varied between 3.2683% (class 3) and 7.7200% (class 2), averaging 5.0718%. Class 2 had the highest standard deviation in percentage errors at 5.4184, while class 5 had the lowest at 2.0670. These metrics highlight variations in pH prediction accuracy across different classes, indicating areas where the model may benefit from further refinement.

Table 4.26 Absolute and relative error of Brix predictions (ANN and RGB values)

Class	class0	class1	class2	class3	class4	class5	Average
Absolute error	0.5825	0.6211	0.7962	0.4156	0.5545	0.5623	0.5742
Standard deviation	0.3530	0.4328	0.5895	0.3327	0.3598	0.3459	0.4074
Relative error	0.0762	0.0646	0.0772	0.0327	0.0387	0.0346	0.0507
Standard deviation	0.0492	0.0449	0.0542	0.0257	0.0263	0.0207	0.0406
Percentage error	7.6241	6.4619	7.7200	3.2683	3.8657	3.4583	5.0718
Standard deviation	4.9245	4.4858	5.4184	2.5706	2.6329	2.0670	4.0584

Table 4.27 compares the average actual and predicted Brix values across different classes. The actual Brix values ranged from 7.8000 for class 0 to 16.1077 for class 5, indicating an increasing trend across classes. The standard deviation in actual Brix values was lowest for class 5 at 0.3639, suggesting less variability, and highest for class 2 at 0.8490. The predicted Brix values closely matched the actual values, with minor deviations noted across all classes. The standard deviation in predicted Brix values was highest for class 2 at 1.2517, indicating more variability in predictions, and lowest for class 5 at 0.2597. Overall, the prediction model was able to effectively capture the trend in Brix values across different classes, although some classes exhibited higher variability in predictions than others.

Table 4.27 Average values and standard deviation of actual and predicted Brix values (ANN and RGB values)

	class0	class1	class2	class3	class4	class5
Average Brix(Actual values)	7.8000	9.5222	10.1077	12.5167	14.4800	16.1077
Standard deviation	0.7711	0.6241	0.8490	0.8248	0.8320	0.3639
Average Brix (Predicted values)	8.0392	9.2589	10.3023	12.5011	14.4563	15.7993
Standard deviation	0.4877	0.8511	1.2517	0.6592	0.3164	0.2597

Figure 4.6 shows the comparison between ANN-predicted Brix values and actual values. The prediction results for class 4 and class 5 mangoes were significantly incorrect.

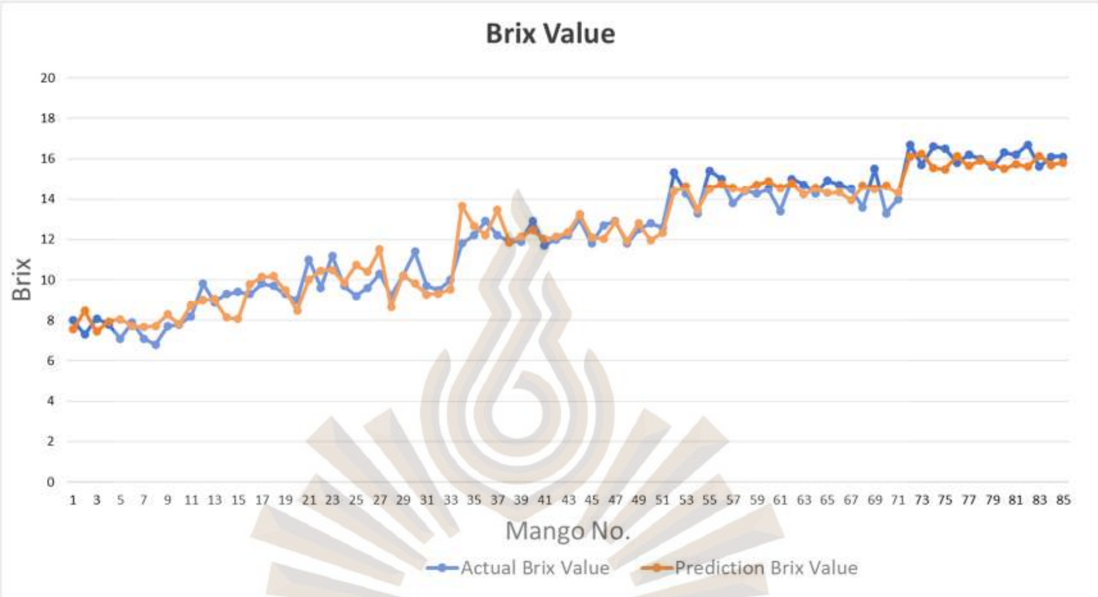

Figure 4.6 Graphs shown the actual and predicted Brix values when predicting using ANN and RGB values.

Source: Researcher, 2024

4.3.3 pH prediction results using linear regression

Table 4.28 shows the errors in predicting pH values based on RGB values measured from three points. The obtained equation was as follows:

$$pH = -2.175 + R1 * 0.01 - G1 * 0.001 - B1 * 0.003 - R2 * 0.004 + G2 * 0.014 - B2 * 0.003 + R3 * 0.001 + G3 * 0.013 + B3 * 0.001 \quad (4.4)$$

The linear regression of RGB and pH had an R^2 of 0.884, indicating rather accurate predictions.

Table 4.28 Results of predicting pH values from RGB values using linear regression

Mango No.	Class	Actual pH Value	Prediction pH Value	Error	Absolute error	Relative error	Percentage error
1	class0	2.08	1.46	-0.62	0.62	0.30	29.81
2	class0	1.97	1.62	-0.35	0.35	0.18	17.77
3	class0	2.02	1.92	-0.10	0.1	0.05	4.95

Table 4.28 Results of predicting pH values from RGB values using linear regression(cont)

Mango No.	Class	Actual pH Value	Prediction pH Value	Error	Absolute error	Relative error	Percentage error
5	class0	1.97	1.71	-0.26	0.26	0.13	13.20
6	class0	1.96	1.82	-0.14	0.14	0.07	7.14
7	class0	2.09	1.72	-0.37	0.37	0.18	17.70
8	class0	1.91	1.76	-0.15	0.15	0.08	7.85
9	class0	1.99	1.87	-0.12	0.12	0.06	6.03
10	class0	1.99	1.85	-0.14	0.14	0.07	7.04
11	class0	2.04	1.77	-0.27	0.27	0.13	13.24
12	class0	2.08	2.06	-0.02	0.02	0.01	0.96
13	class1	2.11	2.06	-0.05	0.05	0.02	2.37
14	class1	2.3	1.88	-0.42	0.42	0.18	18.26
15	class1	2.58	1.73	-0.85	0.85	0.33	32.95
16	class1	2.07	1.86	-0.21	0.21	0.10	10.14
17	class1	2.17	2.21	0.04	0.04	0.02	1.84
18	class1	2.24	2.30	0.06	0.06	0.03	2.68
19	class1	2.01	2.34	0.33	0.33	0.16	16.42
20	class1	2.59	2.13	-0.46	0.46	0.18	17.76
21	class1	2.14	1.78	-0.36	0.36	0.17	16.82
22	class2	2.7	2.38	-0.32	0.32	0.12	11.85
23	class2	2.48	2.48	0.00	0	0.00	0.00
24	class2	2.82	2.45	-0.37	0.37	0.13	13.12
25	class2	2.74	2.38	-0.36	0.36	0.13	13.14
26	class2	2.9	2.54	-0.36	0.36	0.12	12.41
27	class2	2.63	2.34	-0.29	0.29	0.11	11.03
28	class2	2.64	2.61	-0.03	0.03	0.01	1.14
29	class2	2.89	2.09	-0.80	0.8	0.28	27.68
30	class2	2.9	2.48	-0.42	0.42	0.14	14.48
31	class2	2.32	2.28	-0.04	0.04	0.02	1.72
32	class2	2.67	2.23	-0.44	0.44	0.16	16.48
33	class2	2.38	2.20	-0.18	0.18	0.08	7.56
34	class2	2.34	2.35	0.01	0.01	0.00	0.43
35	class3	3.05	2.87	-0.18	0.18	0.06	5.90
36	class3	2.79	2.79	0.00	0	0.00	0.00
37	class3	3.07	2.67	-0.40	0.4	0.13	13.03
38	class3	2.96	2.85	-0.11	0.11	0.04	3.72
39	class3	2.91	2.55	-0.36	0.36	0.12	12.37
40	class3	2.76	2.67	-0.09	0.09	0.03	3.26
41	class3	2.93	2.66	-0.27	0.27	0.09	9.22
42	class3	2.72	2.43	-0.29	0.29	0.11	10.66
43	class3	2.83	2.58	-0.25	0.25	0.09	8.83

Table 4.28 Results of predicting pH values from RGB values using linear regression(cont)

Mango No.	Class	Actual pH Value	Prediction pH Value	Error	Absolute error	Relative error	Percentage error
44	class3	2.73	2.61	-0.12	0.12	0.04	4.40
45	class3	2.75	2.89	0.14	0.14	0.05	5.09
46	class3	2.91	2.62	-0.29	0.29	0.10	9.97
47	class3	2.96	2.54	-0.42	0.42	0.14	14.19
48	class3	2.95	2.76	-0.19	0.19	0.06	6.44
49	class3	2.87	2.37	-0.50	0.5	0.17	17.42
50	class3	2.95	2.68	-0.27	0.27	0.09	9.15
51	class3	2.79	2.51	-0.28	0.28	0.10	10.04
52	class3	2.7	2.71	0.01	0.01	0.00	0.37
53	class4	2.95	2.96	0.01	0.01	0.00	0.34
54	class4	3.35	3.07	-0.28	0.28	0.08	8.36
55	class4	3.25	2.78	-0.47	0.47	0.14	14.46
56	class4	3.17	3.02	-0.15	0.15	0.05	4.73
57	class4	3.39	3.16	-0.23	0.23	0.07	6.78
58	class4	3.37	3.02	-0.35	0.35	0.10	10.39
59	class4	3.44	2.98	-0.46	0.46	0.13	13.37
60	class4	2.88	3.08	0.20	0.2	0.07	6.94
61	class4	3.42	3.09	-0.33	0.33	0.10	9.65
62	class4	3.42	3.03	-0.39	0.39	0.11	11.40
63	class4	3.31	3.15	-0.16	0.16	0.05	4.83
64	class4	2.95	2.96	0.01	0.01	0.00	0.34
65	class4	3.43	3.05	-0.38	0.38	0.11	11.08
66	class4	3.28	2.88	-0.40	0.4	0.12	12.20
67	class4	3.13	2.99	-0.14	0.14	0.04	4.47
68	class4	3.4	2.91	-0.49	0.49	0.14	14.41
69	class4	3.05	3.04	-0.01	0.01	0.00	0.33
70	class4	2.85	3.01	0.16	0.16	0.06	5.61
71	class4	3.37	3.10	-0.27	0.27	0.08	8.01
72	class4	3.15	2.93	-0.22	0.22	0.07	6.98
73	class5	4.14	3.62	-0.52	0.52	0.13	12.56
74	class5	4.23	3.72	-0.51	0.51	0.12	12.06
75	class5	3.83	3.44	-0.39	0.39	0.10	10.18
76	class5	4.15	3.51	-0.64	0.64	0.15	15.42
77	class5	4.25	3.62	-0.63	0.63	0.15	14.82
78	class5	4.13	3.49	-0.64	0.64	0.15	15.50
79	class5	3.96	3.56	-0.40	0.4	0.10	10.10
80	class5	3.82	3.47	-0.35	0.35	0.09	9.16
81	class5	3.83	3.39	-0.44	0.44	0.11	11.49
82	class5	3.88	3.46	-0.42	0.42	0.11	10.82

Table 4.28 Results of predicting pH values from RGB values using linear regression(cont)

Mango No.	Class	Actual pH Value	Prediction pH Value	Error	Absolute error	Relative error	Percentage error
83	class5	4.09	3.44	-0.65	0.65	0.16	15.89
84	class5	3.95	3.67	-0.28	0.28	0.07	7.09
85	class5	3.94	3.65	-0.29	0.29	0.07	7.36

Table 4.29 presents the analysis of absolute and relative errors for different classes. The absolute errors ranged from 0.2467 (class 0) to 0.4738 (class 5), with an average of 0.2918. Standard deviations for absolute errors vary, indicating variability in prediction accuracy across classes, notably high in classes 1 and 2. Relative errors ranged from 0.0773 (class 4) to 0.1325 (class 1), averaging at 0.0997, with corresponding standard deviations reflecting the precision of predictions. Percentage errors showed similar trends. Class 5 exhibited the highest variability. Overall, while predictions were generally accurate, attention to reducing variability, especially in higher error classes, could improve model performance.

Table 4.29 Absolute and relative error of pH predictions (linear regression and RGB values)

Class	class0	class1	class2	class3	class4	class5	Average
Absolute error	0.2467	0.3089	0.2785	0.2317	0.2555	0.4738	0.2918
Standard deviation	0.1698	0.2603	0.2272	0.1388	0.1524	0.1346	0.1895
Relative error	0.1216	0.1325	0.1008	0.0800	0.0773	0.1173	0.0997
Standard deviation	0.0812	0.1016	0.0790	0.0474	0.0445	0.0301	0.0648
Percentage error	12.1564	13.2493	10.0805	8.0029	7.7350	11.7276	9.9692
Standard deviation	8.1201	10.1616	7.9013	4.7364	4.4499	3.0092	6.4833

Table 4.30 compares the average actual and predicted pH values across different classes. The actual average pH values ranged from 2.0150 (class 0) to 4.0154 (class 5), with standard deviations indicating varying degrees of variability within each class. Predicted pH values, on the other hand, ranged from 1.7683 (class 0) to 3.4979 (class 5), generally showing lower values compared to actuals with notable deviations.

Table 4.30 Average and standard deviation of actual and predicted pH values (linear regression and RGB values)

	class0	class1	class2	class3	class4	class5
Average pH (Actual values)	2.0150	2.2456	2.6469	2.8683	3.2280	4.0154
Standard deviation	0.0590	0.2107	0.2101	0.1138	0.1986	0.1559
Average pH (Predicted values)	1.7683	2.0322	2.3700	2.6533	3.0147	3.4979
Standard deviation	0.1544	0.2281	0.1454	0.1453	0.0925	0.1920

Figure 4.7 shows the comparison between the pH values predicted by linear regression and the actual values. The overall prediction results are relatively good.

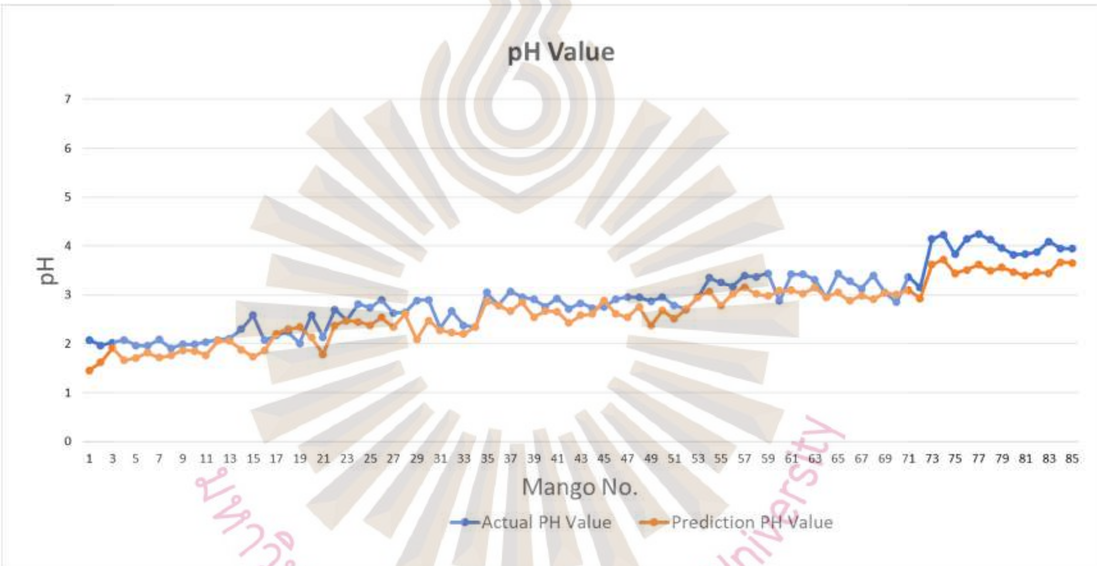

Figure 4.7 Graphs shown the actual and predicted pH values when predicting using linear regression RGB values.

Source: Researcher, 2024

4.3.4 pH prediction results using ANN

Table 4.31 presents the absolute error, relative error, percentage error, and standard deviation of Brix predictions using ANN.

Table 4.31 Results of predicting pH values from RGB values using ANN

Mango No.	Class	Actual pH Value	Prediction pH Value	Error	Absolute error	Relative error	Percentage error
1	class0	1.97	2.16	0.19	0.19	0.10	9.64
2	class0	2.02	2.53	0.51	0.51	0.25	25.25

Table 4.31 Results of predicting pH values from RGB values using ANN(cont)

Mango No.	Class	Actual pH Value	Prediction pH Value	Error	Absolute error	Relative error	Percentage error
3	class0	2.08	2.16	0.08	0.08	0.04	3.85
4	class0	1.97	2.04	0.07	0.07	0.04	3.55
5	class0	1.96	1.83	-0.13	0.13	0.07	6.63
6	class0	2.09	2.17	0.08	0.08	0.04	3.83
7	class0	1.91	1.66	-0.25	0.25	0.13	13.09
8	class0	1.99	2.12	0.13	0.13	0.07	6.53
9	class0	1.99	2.18	0.19	0.19	0.10	9.55
10	class0	2.04	1.77	-0.27	0.27	0.13	13.24
11	class0	2.08	2.06	-0.02	0.02	0.01	0.96
12	class0	2.11	2.33	0.22	0.22	0.10	10.43
13	class1	2.3	2.42	0.12	0.12	0.05	5.22
14	class1	2.58	2.37	-0.21	0.21	0.08	8.14
15	class1	2.07	2.51	0.44	0.44	0.21	21.26
16	class1	2.17	2.43	0.26	0.26	0.12	11.98
17	class1	2.24	1.99	-0.25	0.25	0.11	11.16
18	class1	2.01	2.03	0.02	0.02	0.01	1.00
19	class1	2.59	2.10	-0.49	0.49	0.19	18.92
20	class1	2.14	2.18	0.04	0.04	0.02	1.87
21	class1	2.7	2.67	-0.03	0.03	0.01	1.11
22	class2	2.48	2.88	0.40	0.4	0.16	16.13
23	class2	2.82	2.73	-0.09	0.09	0.03	3.19
24	class2	2.74	2.71	-0.03	0.03	0.01	1.09
25	class2	2.9	2.50	-0.40	0.4	0.14	13.79
26	class2	2.63	2.12	-0.51	0.51	0.19	19.39
27	class2	2.64	1.97	-0.67	0.67	0.25	25.38
28	class2	2.89	2.49	-0.40	0.4	0.14	13.84
29	class2	2.9	2.66	-0.24	0.24	0.08	8.28
30	class2	2.32	2.19	-0.13	0.13	0.06	5.60
31	class2	2.67	2.12	-0.55	0.55	0.21	20.60
32	class2	2.38	2.04	-0.34	0.34	0.14	14.29
33	class2	2.34	2.18	-0.16	0.16	0.07	6.84
34	class2	3.05	3.08	0.03	0.03	0.01	0.98
35	class3	2.79	2.87	0.08	0.08	0.03	2.87
36	class3	3.07	2.77	-0.30	0.3	0.10	9.77
37	class3	2.96	2.55	-0.41	0.41	0.14	13.85
38	class3	2.91	3.11	0.20	0.2	0.07	6.87
39	class3	2.76	2.49	-0.27	0.27	0.10	9.78
40	class3	2.93	3.23	0.30	0.3	0.10	10.24
41	class3	2.72	2.32	-0.40	0.4	0.15	14.71
42	class3	2.83	2.61	-0.22	0.22	0.08	7.77

Table 4.31 Results of predicting pH values from RGB values using ANN(cont)

Mango No.	Class	Actual pH Value	Prediction pH Value	Error	Absolute error	Relative error	Percentage error
43	class3	2.73	2.86	0.13	0.13	0.05	4.76
44	class3	2.75	2.86	0.11	0.11	0.04	4.00
45	class3	2.91	2.83	-0.08	0.08	0.03	2.75
46	class3	2.96	2.66	-0.30	0.3	0.10	10.14
47	class3	2.95	3.04	0.09	0.09	0.03	3.05
48	class3	2.87	2.26	-0.61	0.61	0.21	21.25
49	class3	2.95	2.44	-0.51	0.51	0.17	17.29
50	class3	2.79	2.94	0.15	0.15	0.05	5.38
51	class3	2.7	2.86	0.16	0.16	0.06	5.93
52	class3	2.95	3.19	0.24	0.24	0.08	8.14
53	class4	3.35	3.14	-0.21	0.21	0.06	6.27
54	class4	3.25	3.08	-0.17	0.17	0.05	5.23
55	class4	3.17	3.43	0.26	0.26	0.08	8.20
56	class4	3.39	3.27	-0.12	0.12	0.04	3.54
57	class4	3.37	3.51	0.14	0.14	0.04	4.15
58	class4	3.44	3.46	0.02	0.02	0.01	0.58
59	class4	2.88	3.61	0.73	0.73	0.25	25.35
60	class4	3.42	3.44	0.02	0.02	0.01	0.58
61	class4	3.42	3.33	-0.09	0.09	0.03	2.63
62	class4	3.31	3.25	-0.06	0.06	0.02	1.81
63	class4	2.95	3.02	0.07	0.07	0.02	2.37
64	class4	3.43	3.43	0.00	0	0.00	0.00
65	class4	3.28	3.16	-0.12	0.12	0.04	3.66
66	class4	3.13	3.41	0.28	0.28	0.09	8.95
67	class4	3.4	2.90	-0.50	0.5	0.15	14.71
68	class4	3.05	3.83	0.78	0.78	0.26	25.57
69	class4	2.85	2.94	0.09	0.09	0.03	3.16
70	class4	3.37	3.53	0.16	0.16	0.05	4.75
71	class4	3.15	3.00	-0.15	0.15	0.05	4.76
72	class4	4.14	5.83	1.69	1.69	0.41	40.82
73	class5	4.23	5.60	1.37	1.37	0.32	32.39
74	class5	3.83	4.91	1.08	1.08	0.28	28.20
75	class5	4.15	5.25	1.10	1.1	0.27	26.51
76	class5	4.25	5.33	1.08	1.08	0.25	25.41
77	class5	4.13	4.76	0.63	0.63	0.15	15.25
78	class5	3.96	5.29	1.33	1.33	0.34	33.59
79	class5	3.82	5.19	1.37	1.37	0.36	35.86
80	class5	3.83	5.17	1.34	1.34	0.35	34.99
81	class5	3.88	5.50	1.62	1.62	0.42	41.75
82	class5	4.09	5.02	0.93	0.93	0.23	22.74

Table 4.31 Results of predicting pH values from RGB values using ANN(cont)

Mango No.	Class	Actual pH Value	Prediction pH Value	Error	Absolute error	Relative error	Percentage error
83	class5	3.95	5.36	1.41	1.41	0.36	35.70
84	class5	3.94	4.55	0.61	0.61	0.15	15.48
85	class5	3.94	4.95	1.01	1.01	0.26	25.63

Table 4.32 provides a concise summary of error metrics for Brix prediction across different classes. Absolute errors range from 0.1137 (class 0) to 0.2482 (class 1), with an average of 0.1640. Relative errors vary from 0.0368 (class 5) to 0.1087 (class 1), averaging 0.0591. Percentage errors range from 3.6763% (class 5) to 10.8667% (class 1), with an average of 5.9098%. The standard deviations indicate variability, with class 1 showing the highest variability across all error metrics. These results highlight differences in Brix prediction accuracy, with certain classes showing higher variability and larger errors.

Table 4.32 Absolute and relative error of pH predictions (ANN and RGB values)

Class	class0	class1	class2	class3	class4	class5	Average
Absolute error	0.1137	0.2482	0.1868	0.1471	0.1656	0.1501	0.1640
Standard deviation	0.0524	0.1883	0.1424	0.1092	0.1386	0.1192	0.1293
Relative error	0.0560	0.1087	0.0697	0.0518	0.0528	0.0368	0.0591
Standard deviation	0.0247	0.0858	0.0506	0.0392	0.0488	0.0283	0.0499
Percentage error	5.6014	10.8667	6.9747	5.1776	5.2830	3.6763	5.9098
Standard deviation	2.4652	8.5779	5.0609	3.9175	4.8840	2.8320	4.9912

Table 4.33 shows that predicted pH values are close to actual values, but predictions generally have higher variability. For instance, class 0 has a higher standard deviation in predicted values (0.1351) compared to actual values (0.0625), indicating less consistency in predictions. In contrast, classes 4 and 5 have lower standard deviations in predicted values compared to actual values, suggesting more stable predictions. Overall, while predictions are fairly accurate, improvements are needed for classes with higher variability.

Table 4.33 Average and standard deviation of actual and predicted pH values (ANN and RGB values)

	class0	class1	class2	class3	class4	class5
Average pH (Actual values)	2.0175	2.3111	2.6738	2.8628	3.2875	4.0000
Standard deviation	0.0625	0.2511	0.2380	0.1066	0.2747	0.1524
Average pH (Predicted values)	2.0526	2.3165	2.6311	2.8992	3.3207	3.8622
Standard deviation	0.1351	0.2251	0.2189	0.1686	0.0993	0.0999

Figure 4.8 shows the comparison between the pH values predicted by ANN and the actual values.

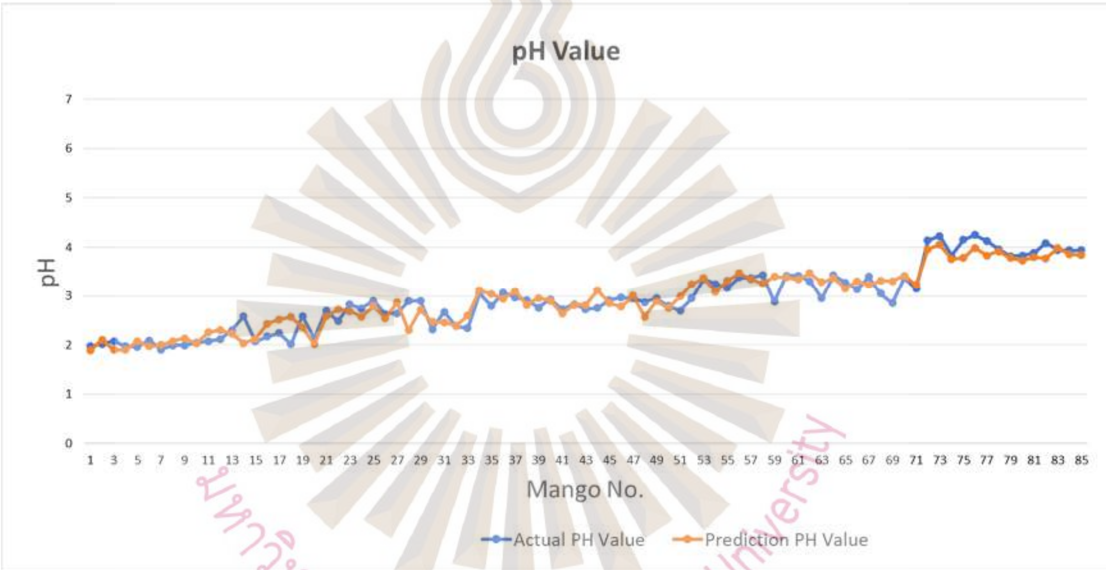

Figure 4.8 Graphs shown the actual and predicted pH values when predicting using ANN and RGB values.

Source: Researcher, 2024

The above tables 4.28, 4.29, 4.30, 4.31, 4.32, and 4.33 show the relationship between image RGB values and mango juice pH values. Tables 4.29 and 4.31 show the error values between the predicted values and the relative values. On average, the results of ANN are superior to linear regression.

4.4 Brix and pH prediction results using Lab

4.4.1 Brix prediction results using linear regression

Table 4.34 shows the errors in predicting pH values based on Lab values measured from three points.The obtained equation was as follows:

$$Brix = -2.769 + L1 * 0.047 + a1 * 0.117 - b1 * 0.136 + L2 * 0.085 - a2 * 0.035 + b2 * 0.054 + L3 * 0.117 + a3 * 0.113 + b3 * 0.036 \quad (4.5)$$

The Lab and Brix linear regression had an R^2 of 0.930, indicating good predictions.

Table 4.34 presents the analysis of absolute and relative errors.

Table 4.34 Results of predicting Brix values from Lab values using linear regression

Mango No.	Class	Actual Brix Value	Prediction Brix Value	Error	Absolute error	Relative error	Percentage error
1	class0	8	7.17	-0.83	0.83	0.10	10.38
2	class0	7.3	7.21	-0.09	0.09	0.01	1.23
3	class0	8.1	8.54	0.44	0.44	0.05	5.43
4	class0	7.8	7.92	0.12	0.12	0.02	1.54
5	class0	7.1	8.10	1.00	1	0.14	14.08
6	class0	7.9	7.27	-0.63	0.63	0.08	7.97
7	class0	7.1	9.00	1.90	1.9	0.27	26.76
8	class0	6.8	7.92	1.12	1.12	0.16	16.47
9	class0	7.7	8.10	0.40	0.4	0.05	5.19
10	class0	7.8	7.89	0.09	0.09	0.01	1.15
11	class0	8.2	7.91	-0.29	0.29	0.04	3.54
12	class0	9.8	9.19	-0.61	0.61	0.06	6.22
13	class1	8.9	8.52	-0.38	0.38	0.04	4.27
14	class1	9.3	8.66	-0.64	0.64	0.07	6.88
15	class1	9.4	7.98	-1.42	1.42	0.15	15.11
16	class1	9.3	7.71	-1.59	1.59	0.17	17.10
17	class1	9.8	9.58	-0.22	0.22	0.02	2.24
18	class1	9.7	9.68	-0.02	0.02	0.00	0.21
19	class1	9.3	10.01	0.71	0.71	0.08	7.63
20	class1	9	9.95	0.95	0.95	0.11	10.56
21	class1	11	9.05	-1.95	1.95	0.18	17.73
22	class2	9.6	10.38	0.78	0.78	0.08	8.13
23	class2	11.2	10.07	-1.13	1.13	0.10	10.09
24	class2	9.7	10.24	0.54	0.54	0.06	5.57
25	class2	9.2	10.24	1.04	1.04	0.11	11.30
26	class2	9.6	10.40	0.80	0.8	0.08	8.33

Table 4.34 Results of predicting Brix values from Lab values using linear regression(cont)

Mango No.	Class	Actual Brix Value	Prediction Brix Value	Error	Absolute error	Relative error	Percentage error
27	class2	10.3	10.41	0.11	0.11	0.01	1.07
28	class2	9.2	11.06	1.86	1.86	0.20	20.22
29	class2	10.2	9.14	-1.06	1.06	0.10	10.39
30	class2	11.4	9.97	-1.43	1.43	0.13	12.54
31	class2	9.7	10.10	0.40	0.4	0.04	4.12
32	class2	9.5	9.54	0.04	0.04	0.00	0.42
33	class2	10	9.29	-0.71	0.71	0.07	7.10
34	class2	11.8	9.78	-2.02	2.02	0.17	17.12
35	class3	12.2	14.29	2.09	2.09	0.17	17.13
36	class3	12.9	12.89	-0.01	0.01	0.00	0.08
37	class3	12.2	11.94	-0.26	0.26	0.02	2.13
38	class3	11.9	13.70	1.80	1.8	0.15	15.13
39	class3	11.9	11.59	-0.31	0.31	0.03	2.61
40	class3	12.9	12.01	-0.89	0.89	0.07	6.90
41	class3	11.7	12.07	0.37	0.37	0.03	3.16
42	class3	12	11.73	-0.27	0.27	0.02	2.25
43	class3	12.2	12.27	0.07	0.07	0.01	0.57
44	class3	13	11.81	-1.19	1.19	0.09	9.15
45	class3	11.8	12.81	1.01	1.01	0.09	8.56
46	class3	12.7	12.13	-0.57	0.57	0.04	4.49
47	class3	12.9	12.60	-0.30	0.3	0.02	2.33
48	class3	11.8	12.55	0.75	0.75	0.06	6.36
49	class3	12.5	12.57	0.07	0.07	0.01	0.56
50	class3	12.8	13.11	0.31	0.31	0.02	2.42
51	class3	12.6	12.09	-0.51	0.51	0.04	4.05
52	class3	15.3	12.76	-2.54	2.54	0.17	16.60
53	class4	14.3	14.84	0.54	0.54	0.04	3.78
54	class4	13.3	15.37	2.07	2.07	0.16	15.56
55	class4	15.4	14.31	-1.09	1.09	0.07	7.08
56	class4	15	14.71	-0.29	0.29	0.02	1.93
57	class4	13.8	14.37	0.57	0.57	0.04	4.13
58	class4	14.4	14.93	0.53	0.53	0.04	3.68
59	class4	14.3	14.14	-0.16	0.16	0.01	1.12
60	class4	14.5	14.17	-0.33	0.33	0.02	2.28
61	class4	13.4	14.15	0.75	0.75	0.06	5.60
62	class4	15	14.72	-0.28	0.28	0.02	1.87
63	class4	14.7	14.03	-0.67	0.67	0.05	4.56
64	class4	14.3	14.28	-0.02	0.02	0.00	0.14
65	class4	14.9	14.53	-0.37	0.37	0.02	2.48

Table 4.34 Results of predicting Brix values from Lab values using linear regression(cont)

Mango No.	Class	Actual Brix Value	Prediction Brix Value	Error	Absolute error	Relative error	Percentage error
66	class4	14.7	14.67	-0.03	0.03	0.00	0.20
67	class4	14.5	14.66	0.16	0.16	0.01	1.10
68	class4	13.6	14.27	0.67	0.67	0.05	4.93
69	class4	15.5	14.34	-1.16	1.16	0.07	7.48
70	class4	13.3	14.09	0.79	0.79	0.06	5.94
71	class4	14	14.19	0.19	0.19	0.01	1.36
72	class4	16.7	14.48	-2.22	2.22	0.13	13.29
73	class5	15.7	15.58	-0.12	0.12	0.01	0.76
74	class5	16.6	16.20	-0.40	0.4	0.02	2.41
75	class5	16.5	16.32	-0.18	0.18	0.01	1.09
76	class5	15.8	15.74	-0.06	0.06	0.00	0.38
77	class5	16.2	16.01	-0.19	0.19	0.01	1.17
78	class5	16	15.88	-0.12	0.12	0.01	0.75
79	class5	15.6	15.36	-0.24	0.24	0.02	1.54
80	class5	16.3	15.90	-0.40	0.4	0.02	2.45
81	class5	16.2	15.17	-1.03	1.03	0.06	6.36
82	class5	16.7	15.25	-1.45	1.45	0.09	8.68
83	class5	15.6	15.99	0.39	0.39	0.03	2.50
84	class5	16.1	16.04	-0.06	0.06	0.00	0.37
85	class5	16.1	16.11	0.01	0.01	0.00	0.07

Table 4.35 summarizes the performance metrics for pH prediction across different classes. Absolute errors ranged from 0.3577 to 0.9169, with an average error of 0.6845. The error standard deviations varied across classes, indicating different levels of variability in prediction accuracy. Relative errors were relatively low, ranging from 0.0220 to 0.0908, with an average of 0.0611, suggesting generally accurate predictions relative to the actual pH values. Percentage errors ranged from 2.1953% to 9.0803%, averaging 6.1132%, reflecting the accuracy of pH predictions across the classes.

Table 4.35 Absolute and relative error of Brix predictions (linear regression and Lab values)

Class	class0	class1	class2	class3	class4	class5	Average
Absolute error	0.6267	0.8756	0.9169	0.7400	0.6445	0.3577	0.6845
Standard deviation	0.5290	0.6574	0.6024	0.7336	0.6034	0.4217	0.6137
Relative error	0.0833	0.0908	0.0895	0.0580	0.0443	0.0220	0.0611
Standard deviation	0.0761	0.0645	0.0570	0.0548	0.0406	0.0254	0.0574
Percentage error	8.3315	9.0803	8.9541	5.8039	4.4255	2.1953	6.1132
Standard deviation	7.6098	6.4524	5.6980	5.4849	4.0576	2.5431	5.7352

Table 4.36 compares actual and predicted average Brix values, along with their respective standard deviations across different classes. The actual average Brix values range from 7.8000 to 16.1077, with corresponding standard deviations indicating variability within each class. Predicted average Brix values closely approximate the actual values, ranging from 8.0183 to 15.7165. The standard deviations of predicted values reflected varying levels of error across classes.

Table 4.36 Average and standard deviation of actual and predicted Brix values (linear regression and Lab values)

	class0	class1	class2	class3	class4	class5
Average Brix(Actual values)	7.8000	9.5222	10.1077	12.5167	14.4800	16.1077
Standard deviation	0.7711	0.6241	0.8490	0.8248	0.8320	0.3639
Average Brix (Predicted values)	8.0183	9.0156	10.0477	12.4956	14.4616	15.7165
Standard deviation	0.6469	0.8494	0.5165	0.7005	0.3499	0.5016

Figure 4.9 shows the comparison between the Brix values predicted by linear regression and the actual values. It could be seen that the predicted results had a similar range of changes to the actual results, but there were still some errors.

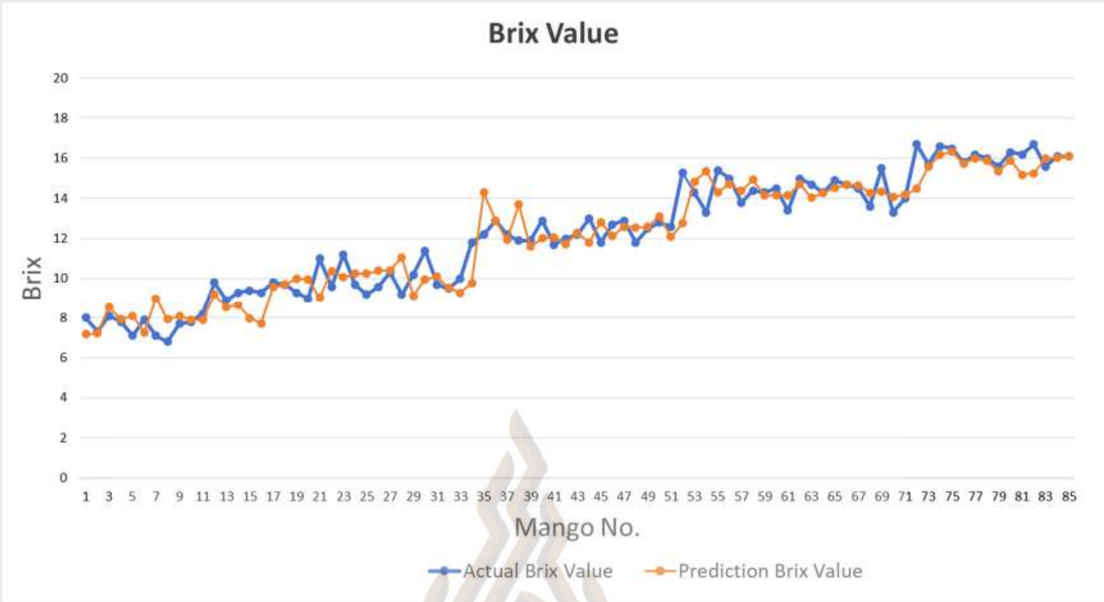

Figure 4.9 Graphs shown the actual and predicted Brix values when predicting using linear regression and Lab values.
Source: Researcher, 2024

4.4.2 Brix prediction results using ANN

Table 4.37 shows the results predicting Brix using ANN and three Lab points.

Table 4.37 Results of predicting Brix values from Lab values using ANN

Mango No.	Class	Actual Brix Value	Prediction Brix Value	Error	Absolute error	Relative error	Percentage error
1	class0	8	7.22	-0.78	0.78	0.10	9.75
2	class0	7.3	6.55	-0.75	0.75	0.10	10.27
3	class0	8.1	6.97	-1.13	1.13	0.14	13.95
4	class0	7.8	6.76	-1.04	1.04	0.13	13.33
5	class0	7.1	7.99	0.89	0.89	0.13	12.54
6	class0	7.9	6.67	-1.23	1.23	0.16	15.57
7	class0	7.1	7.57	0.47	0.47	0.07	6.62
8	class0	6.8	7.51	0.71	0.71	0.10	10.44
9	class0	7.7	7.77	0.07	0.07	0.01	0.91
10	class0	7.8	7.48	-0.32	0.32	0.04	4.10
11	class0	8.2	7.58	-0.62	0.62	0.08	7.56
12	class0	9.8	11.23	1.43	1.43	0.15	14.59
13	class1	8.9	10.76	1.86	1.86	0.21	20.90
14	class1	9.3	9.57	0.27	0.27	0.03	2.90
15	class1	9.4	10.29	0.89	0.89	0.09	9.47
16	class1	9.3	7.92	-1.38	1.38	0.15	14.84

Table 4.37 Results of predicting Brix values from Lab values using ANN(cont)

Mango No.	Class	Actual Brix Value	Prediction Brix Value	Error	Absolute error	Relative error	Percentage error
17	class1	9.8	8.49	-1.31	1.31	0.13	13.37
18	class1	9.7	9.05	-0.65	0.65	0.07	6.70
19	class1	9.3	7.73	-1.57	1.57	0.17	16.88
20	class1	9	8.03	-0.97	0.97	0.11	10.78
21	class1	11	8.80	-2.20	2.2	0.20	20.00
22	class2	9.6	11.38	1.78	1.78	0.19	18.54
23	class2	11.2	8.90	-2.30	2.3	0.21	20.54
24	class2	9.7	10.37	0.67	0.67	0.07	6.91
25	class2	9.2	8.80	-0.40	0.4	0.04	4.35
26	class2	9.6	6.64	-2.96	2.96	0.31	30.83
27	class2	10.3	7.74	-2.56	2.56	0.25	24.85
28	class2	9.2	9.04	-0.16	0.16	0.02	1.74
29	class2	10.2	9.00	-1.20	1.2	0.12	11.76
30	class2	11.4	7.97	-3.43	3.43	0.30	30.09
31	class2	9.7	8.05	-1.65	1.65	0.17	17.01
32	class2	9.5	8.41	-1.09	1.09	0.11	11.47
33	class2	10	7.92	-2.08	2.08	0.21	20.80
34	class2	11.8	14.67	2.87	2.87	0.24	24.32
35	class3	12.2	12.34	0.14	0.14	0.01	1.15
36	class3	12.9	12.23	-0.67	0.67	0.05	5.19
37	class3	12.2	13.94	1.74	1.74	0.14	14.26
38	class3	11.9	13.55	1.65	1.65	0.14	13.87
39	class3	11.9	12.09	0.19	0.19	0.02	1.60
40	class3	12.9	12.79	-0.11	0.11	0.01	0.85
41	class3	11.7	12.26	0.56	0.56	0.05	4.79
42	class3	12	13.48	1.48	1.48	0.12	12.33
43	class3	12.2	11.69	-0.51	0.51	0.04	4.18
44	class3	13	12.89	-0.11	0.11	0.01	0.85
45	class3	11.8	12.19	0.39	0.39	0.03	3.31
46	class3	12.7	10.47	-2.23	2.23	0.18	17.56
47	class3	12.9	11.30	-1.60	1.6	0.12	12.40
48	class3	11.8	13.78	1.98	1.98	0.17	16.78
49	class3	12.5	12.80	0.30	0.3	0.02	2.40
50	class3	12.8	12.34	-0.46	0.46	0.04	3.59
51	class3	12.6	11.69	-0.91	0.91	0.07	7.22
52	class3	15.3	14.44	-0.86	0.86	0.06	5.62
53	class4	14.3	14.75	0.45	0.45	0.03	3.15
54	class4	13.3	15.48	2.18	2.18	0.16	16.39
55	class4	15.4	13.55	-1.85	1.85	0.12	12.01
56	class4	15	14.92	-0.08	0.08	0.01	0.53

Table 4.37 Results of predicting Brix values from Lab values using ANN(cont)

Mango No.	Class	Actual Brix Value	Prediction Brix Value	Error	Absolute error	Relative error	Percentage error
57	class4	13.8	15.48	1.68	1.68	0.12	12.17
58	class4	14.4	14.99	0.59	0.59	0.04	4.10
59	class4	14.3	12.05	-2.25	2.25	0.16	15.73
60	class4	14.5	13.74	-0.76	0.76	0.05	5.24
61	class4	13.4	15.16	1.76	1.76	0.13	13.13
62	class4	15	14.13	-0.87	0.87	0.06	5.80
63	class4	14.7	14.53	-0.17	0.17	0.01	1.16
64	class4	14.3	13.48	-0.82	0.82	0.06	5.73
65	class4	14.9	13.89	-1.01	1.01	0.07	6.78
66	class4	14.7	15.26	0.56	0.56	0.04	3.81
67	class4	14.5	13.28	-1.22	1.22	0.08	8.41
68	class4	13.6	13.37	-0.23	0.23	0.02	1.69
69	class4	15.5	12.47	-3.03	3.03	0.20	19.55
70	class4	13.3	13.85	0.55	0.55	0.04	4.14
71	class4	14	13.89	-0.11	0.11	0.01	0.79
72	class4	16.7	16.71	0.01	0.01	0.00	0.06
73	class5	15.7	18.82	3.12	3.12	0.20	19.87
74	class5	16.6	19.87	3.27	3.27	0.20	19.70
75	class5	16.5	17.26	0.76	0.76	0.05	4.61
76	class5	15.8	18.44	2.64	2.64	0.17	16.71
77	class5	16.2	17.91	1.71	1.71	0.11	10.56
78	class5	16	16.62	0.62	0.62	0.04	3.88
79	class5	15.6	17.34	1.74	1.74	0.11	11.15
80	class5	16.3	15.15	-1.15	1.15	0.07	7.06
81	class5	16.2	16.18	-0.02	0.02	0.00	0.12
82	class5	16.7	18.39	1.69	1.69	0.10	10.12
83	class5	15.6	17.49	1.89	1.89	0.12	12.12
84	class5	16.1	17.91	1.81	1.81	0.11	11.24
85	class5	16.1	18.64	2.54	2.54	0.16	15.78

Table 4.38 shows error metrics for pH predictions across different classes. Absolute errors ranged from 0.4287 to 0.7981, with class 2 having the highest variability. Relative errors were between 0.0324 and 0.0767, with class 5 showing the lowest variability. Percentage errors varied from 3.2398% to 7.6728%, with class 5 exhibiting the lowest variability. Overall, while the model's predictions were reasonably accurate, there was notable variability in errors, particularly in percentage errors, indicating areas for potential improvement.

Table 4.38 Absolute and relative error of Brix predictions (ANN and Lab values)

Class	class0	class1	class2	class3	class4	class5	Average
Absolute error	0.4738	0.4287	0.7981	0.5048	0.6176	0.5256	0.5670
Standard deviation	0.3191	0.4497	0.5177	0.3219	0.4398	0.2814	0.4002
Relative error	0.0617	0.0449	0.0767	0.0398	0.0431	0.0324	0.0487
Standard deviation	0.0442	0.0478	0.0443	0.0244	0.0310	0.0170	0.0367
Percentage error	6.1699	4.4864	7.6728	3.9818	4.3058	3.2398	4.8714
Standard deviation	4.4249	4.7810	4.4295	2.4437	3.1005	1.6996	3.6699

Table 4.39 shows that predicted Brix values are generally close to actual values, with averages ranging from 7.8000 to 16.1077. While the predictions aligned well with actual values, there was greater variability in some classes, particularly class 2, which had the highest standard deviation in predictions. Overall, the model could capture Brix trends effectively.

Table 4.39 Average and standard deviation of actual and predicted Brix values (ANN and Lab values)

	class0	class1	class2	class3	class4	class5
Average Brix(Actual values)	7.8000	9.5222	10.1077	12.5167	14.4800	16.1077
Standard deviation	0.7711	0.6241	0.8490	0.8248	0.8320	0.3639
Average Brix (Predicted values)	8.0235	9.3090	10.2660	12.4795	14.5355	15.7261
Standard deviation	0.4923	0.7878	1.2568	0.6286	0.3275	0.2535

Tables 4.34, 4.35, 4.36, 4.37, 4.38 and 4.39 above show the relationship between Lab values and sugar content values predicted by linear regression equations and ANN predictors. Figure 4.10 shows the comparison between the Brix values predicted by ANN and the actual values. The predicted values of the mango Brix still deviated from the actual values.

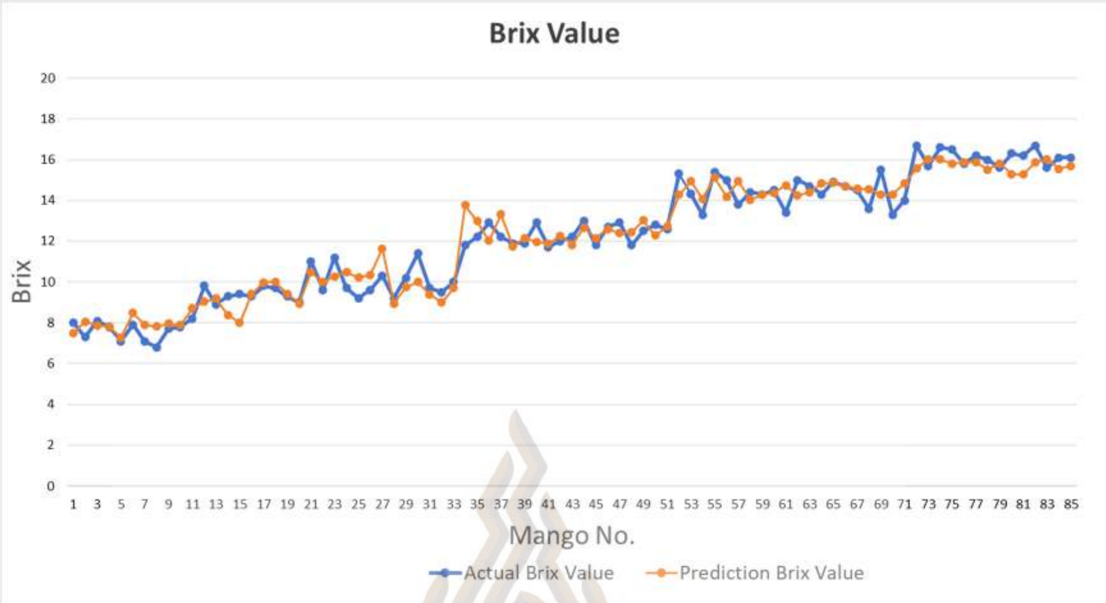

Figure 4.10 Graphs shown the actual and predicted Brix values when predicting using ANN and Lab values.

Source: Researcher, 2024

The comparison between linear regression equation prediction and ANN prediction showed that ANN achieved better results.

4.4.3 pH prediction results using linear regression

When using linear regression, Table 4.40 shows the errors in predicting pH values based on Lab values measured from three points. After applying linear regression to predict pH values, the obtained equation, which had R^2 of 0.867, was as follows:

$$pH = -0.779 + L1 * 0.015 + a1 * 0.057 - b1 * 0.025 + L2 * 0.025 - a2 * 0.039 - b2 * 0.054 + L3 * 0.3 + a3 * 0.018 + b3 * 0.013 \quad (4.6)$$

Table 4.40 Results of predicting pH values from Lab values using linear regression

Mango No.	Class	Actual pH Value	Prediction pH Value	Error	Absolute error	Relative error	Percentage error
1	class0	1.97	1.89	-0.08	0.08	0.04	4.06
2	class0	2.02	1.89	-0.13	0.13	0.06	6.44
3	class0	2.08	2.37	0.29	0.29	0.14	13.94
4	class0	1.97	2.06	0.09	0.09	0.05	4.57
5	class0	1.96	2.14	0.18	0.18	0.09	9.18
6	class0	2.09	1.82	-0.27	0.27	0.13	12.92
7	class0	1.91	2.26	0.35	0.35	0.18	18.32

Table 4.40 Results of predicting pH values from Lab values using linear regression(cont)

Mango No.	Class	Actual pH Value	Prediction pH Value	Error	Absolute error	Relative error	Percentage error
8	class0	1.99	2.03	0.04	0.04	0.02	2.01
9	class0	1.99	2.11	0.12	0.12	0.06	6.03
10	class0	2.04	2.19	0.15	0.15	0.07	7.35
11	class0	2.08	2.03	-0.05	0.05	0.02	2.40
12	class0	2.11	2.27	0.16	0.16	0.08	7.58
13	class1	2.3	2.11	-0.19	0.19	0.08	8.26
14	class1	2.58	2.10	-0.48	0.48	0.19	18.60
15	class1	2.07	1.98	-0.09	0.09	0.04	4.35
16	class1	2.17	2.05	-0.12	0.12	0.06	5.53
17	class1	2.24	2.47	0.23	0.23	0.10	10.27
18	class1	2.01	2.35	0.34	0.34	0.17	16.92
19	class1	2.59	2.52	-0.07	0.07	0.03	2.70
20	class1	2.14	2.47	0.33	0.33	0.15	15.42
21	class1	2.7	2.25	-0.45	0.45	0.17	16.67
22	class2	2.48	2.57	0.09	0.09	0.04	3.63
23	class2	2.82	2.66	-0.16	0.16	0.06	5.67
24	class2	2.74	2.68	-0.06	0.06	0.02	2.19
25	class2	2.9	2.73	-0.17	0.17	0.06	5.86
26	class2	2.63	2.90	0.27	0.27	0.10	10.27
27	class2	2.64	2.59	-0.05	0.05	0.02	1.89
28	class2	2.89	2.75	-0.14	0.14	0.05	4.84
29	class2	2.9	2.39	-0.51	0.51	0.18	17.59
30	class2	2.32	2.57	0.25	0.25	0.11	10.78
31	class2	2.67	2.60	-0.07	0.07	0.03	2.62
32	class2	2.38	2.50	0.12	0.12	0.05	5.04
33	class2	2.34	2.53	0.19	0.19	0.08	8.12
34	class2	3.05	2.59	-0.46	0.46	0.15	15.08
35	class3	2.79	3.30	0.51	0.51	0.18	18.28
36	class3	3.07	3.15	0.08	0.08	0.03	2.61
37	class3	2.96	2.91	-0.05	0.05	0.02	1.69
38	class3	2.91	3.10	0.19	0.19	0.07	6.53
39	class3	2.76	2.62	-0.14	0.14	0.05	5.07
40	class3	2.93	2.75	-0.18	0.18	0.06	6.14
41	class3	2.72	2.80	0.08	0.08	0.03	2.94
42	class3	2.83	2.71	-0.12	0.12	0.04	4.24
43	class3	2.73	2.88	0.15	0.15	0.05	5.49
44	class3	2.75	2.64	-0.11	0.11	0.04	4.00
45	class3	2.91	2.95	0.04	0.04	0.01	1.37
46	class3	2.96	2.79	-0.17	0.17	0.06	5.74
47	class3	2.95	2.88	-0.07	0.07	0.02	2.37

Table 4.40 Results of predicting pH values from Lab values using linear regression(cont)

Mango No.	Class	Actual pH Value	Prediction pH Value	Error	Absolute error	Relative error	Percentage error
48	class3	2.87	2.91	0.04	0.04	0.01	1.39
49	class3	2.95	3.00	0.05	0.05	0.02	1.69
50	class3	2.79	2.96	0.17	0.17	0.06	6.09
51	class3	2.7	2.89	0.19	0.19	0.07	7.04
52	class3	2.95	3.02	0.07	0.07	0.02	2.37
53	class4	3.35	3.34	-0.01	0.01	0.00	0.30
54	class4	3.25	3.65	0.40	0.4	0.12	12.31
55	class4	3.17	3.40	0.23	0.23	0.07	7.26
56	class4	3.39	3.55	0.16	0.16	0.05	4.72
57	class4	3.37	3.33	-0.04	0.04	0.01	1.19
58	class4	3.44	3.49	0.05	0.05	0.01	1.45
59	class4	2.88	3.21	0.33	0.33	0.11	11.46
60	class4	3.42	3.35	-0.07	0.07	0.02	2.05
61	class4	3.42	3.36	-0.06	0.06	0.02	1.75
62	class4	3.31	3.45	0.14	0.14	0.04	4.23
63	class4	2.95	3.22	0.27	0.27	0.09	9.15
64	class4	3.43	3.33	-0.10	0.1	0.03	2.92
65	class4	3.28	3.51	0.23	0.23	0.07	7.01
66	class4	3.13	3.47	0.34	0.34	0.11	10.86
67	class4	3.4	3.40	0.00	0	0.00	0.00
68	class4	3.05	3.41	0.36	0.36	0.12	11.80
69	class4	2.85	3.43	0.58	0.58	0.20	20.35
70	class4	3.37	3.40	0.03	0.03	0.01	0.89
71	class4	3.15	3.35	0.20	0.2	0.06	6.35
72	class4	4.14	3.43	-0.71	0.71	0.17	17.15
73	class5	4.23	3.83	-0.40	0.4	0.09	9.46
74	class5	3.83	3.91	0.08	0.08	0.02	2.09
75	class5	4.15	3.96	-0.19	0.19	0.05	4.58
76	class5	4.25	3.81	-0.44	0.44	0.10	10.35
77	class5	4.13	3.81	-0.32	0.32	0.08	7.75
78	class5	3.96	3.89	-0.07	0.07	0.02	1.77
79	class5	3.82	3.70	-0.12	0.12	0.03	3.14
80	class5	3.83	3.85	0.02	0.02	0.01	0.52
81	class5	3.88	3.65	-0.23	0.23	0.06	5.93
82	class5	4.09	3.64	-0.45	0.45	0.11	11.00
83	class5	3.95	3.81	-0.14	0.14	0.04	3.54
84	class5	3.94	3.90	-0.04	0.04	0.01	1.02
85	class5	3.94	3.92	-0.02	0.02	0.01	0.51

Table 4.41 provides a summary of errors and deviations for different classes. The average absolute error ranged from 0.1592 to 0.2556, with an overall average of 0.1881. Standard deviations for absolute errors varied between 0.0981 and 0.1938. Relative errors ranged from 0.0720 to 0.1097, with a mean of 0.0667. The corresponding standard deviations for relative errors ranged from 0.0380 to 0.0608. Percentage errors showed values between 4.7266% and 10.9685%, with an average of 6.6711%. Standard deviations for percentage errors ranged from 3.7956 to 6.0779%.

Table 4.41 Absolute and relative error of pH predictions (linear regression and Lab values)

Class	class0	class1	class2	class3	class4	class5	Average
Absolute error	0.1592	0.2556	0.1954	0.1339	0.2155	0.1938	0.1881
Standard deviation	0.0981	0.1525	0.1456	0.1083	0.1938	0.1599	0.1498
Relative error	0.0790	0.1097	0.0720	0.0473	0.0666	0.0474	0.0667
Standard deviation	0.0494	0.0608	0.0495	0.0390	0.0584	0.0380	0.0517
Percentage error	7.9012	10.9685	7.1990	4.7266	6.6599	4.7425	6.6711
Standard deviation	4.9433	6.0779	4.9525	3.8970	5.8378	3.7956	5.1658

Table 4.42 summarizes pH values for actual and predicted data across different classes. The average pH values for actual measurements ranged from 2.0175 to 4.0000, with corresponding standard deviations ranging from 0.0625 to 0.2747. Predicted average pH values ranged from 2.0883 to 3.7936, with standard deviations between 0.1059 and 0.1750.

Table 4.42 Average and standard deviation of actual and predicted pH values (linear regression and Lab values)

	class0	class1	class2	class3	class4	class5
Average pH (Actual values)	2.0175	2.3111	2.6738	2.8628	3.2875	4.0000
Standard deviation	0.0625	0.2511	0.2380	0.1066	0.2747	0.1524
Average pH (Predicted values)	2.0883	2.2556	2.6200	2.9033	3.4026	3.7936
Standard deviation	0.1688	0.2045	0.1275	0.1750	0.1059	0.1437

Figure 4.11 shows the comparison between the pH value predicted by linear regression and the actual value.

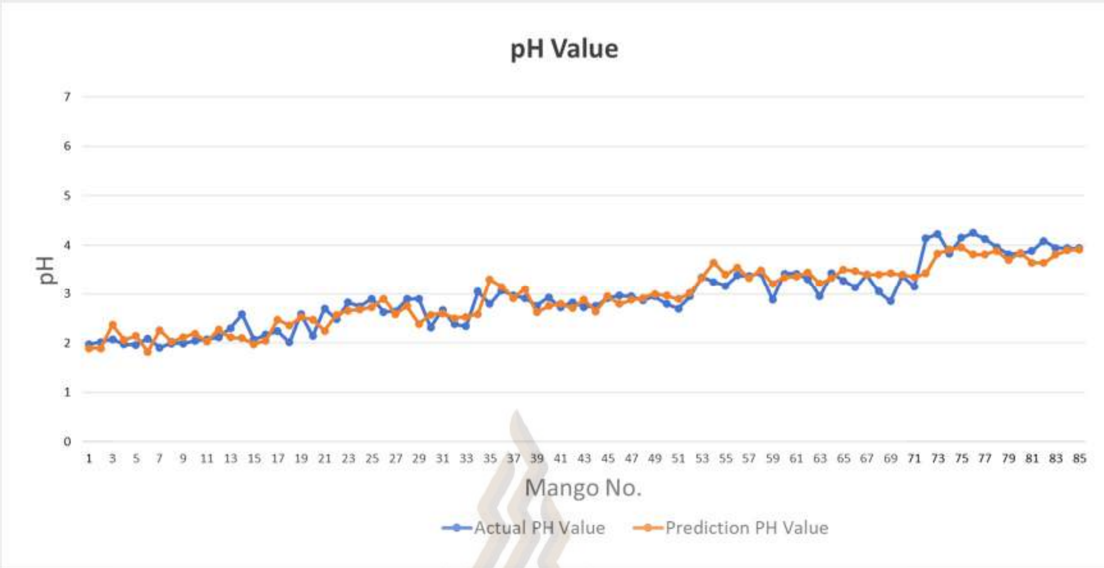

Figure 4.11Graphs shown the actual and predicted Brix values when predicting using linear regression and Lab values.

Source: Researcher, 2024

4.4.4 pH prediction results using ANN

Table 4.43 shows the errors in predicting pH values based on Lab values.

Table 4.43 Results of predicting pH values from Lab values using ANN

Mango No.	Class	Actual pH Value	Prediction pH Value	Error	Absolute error	Relative error	Percentage error
1	class0	1.97	1.96	-0.01	0.01	0.01	0.51
2	class0	2.02	2.06	0.04	0.04	0.02	1.98
3	class0	2.08	2.02	-0.06	0.06	0.03	2.88
4	class0	1.97	2.09	0.12	0.12	0.06	6.09
5	class0	1.96	2.14	0.18	0.18	0.09	9.18
6	class0	2.09	1.98	-0.11	0.11	0.05	5.26
7	class0	1.91	1.98	0.07	0.07	0.04	3.66
8	class0	1.99	1.94	-0.05	0.05	0.03	2.51
9	class0	1.99	2.03	0.04	0.04	0.02	2.01
10	class0	2.04	1.95	-0.09	0.09	0.04	4.41
11	class0	2.08	2.10	0.02	0.02	0.01	0.96
12	class0	2.11	2.58	0.47	0.47	0.22	22.27
13	class1	2.3	2.44	0.14	0.14	0.06	6.09
14	class1	2.58	2.41	-0.17	0.17	0.07	6.59
15	class1	2.07	2.55	0.48	0.48	0.23	23.19
16	class1	2.17	2.12	-0.05	0.05	0.02	2.30

Table 4.43 Results of predicting pH values from Lab values using ANN(cont)

Mango No.	Class	Actual pH Value	Prediction pH Value	Error	Absolute error	Relative error	Percentage error
17	class1	2.24	2.12	-0.12	0.12	0.05	5.36
18	class1	2.01	2.43	0.42	0.42	0.21	20.90
19	class1	2.59	2.34	-0.25	0.25	0.10	9.65
20	class1	2.14	2.14	0.00	0	0.00	0.00
21	class1	2.7	2.49	-0.21	0.21	0.08	7.78
22	class2	2.48	2.98	0.50	0.5	0.20	20.16
23	class2	2.82	2.88	0.06	0.06	0.02	2.13
24	class2	2.74	2.95	0.21	0.21	0.08	7.66
25	class2	2.9	2.83	-0.07	0.07	0.02	2.41
26	class2	2.63	2.02	-0.61	0.61	0.23	23.19
27	class2	2.64	2.23	-0.41	0.41	0.16	15.53
28	class2	2.89	2.71	-0.18	0.18	0.06	6.23
29	class2	2.9	2.88	-0.02	0.02	0.01	0.69
30	class2	2.32	2.41	0.09	0.09	0.04	3.88
31	class2	2.67	2.43	-0.24	0.24	0.09	8.99
32	class2	2.38	2.43	0.05	0.05	0.02	2.10
33	class2	2.34	2.53	0.19	0.19	0.08	8.12
34	class2	3.05	3.26	0.21	0.21	0.07	6.89
35	class3	2.79	2.90	0.11	0.11	0.04	3.94
36	class3	3.07	2.90	-0.17	0.17	0.06	5.54
37	class3	2.96	2.84	-0.12	0.12	0.04	4.05
38	class3	2.91	3.04	0.13	0.13	0.04	4.47
39	class3	2.76	2.76	0.00	0	0.00	0.00
40	class3	2.93	2.92	-0.01	0.01	0.00	0.34
41	class3	2.72	2.83	0.11	0.11	0.04	4.04
42	class3	2.83	3.06	0.23	0.23	0.08	8.13
43	class3	2.73	2.84	0.11	0.11	0.04	4.03
44	class3	2.75	3.08	0.33	0.33	0.12	12.00
45	class3	2.91	2.81	-0.10	0.1	0.03	3.44
46	class3	2.96	2.55	-0.41	0.41	0.14	13.85
47	class3	2.95	2.92	-0.03	0.03	0.01	1.02
48	class3	2.87	3.05	0.18	0.18	0.06	6.27
49	class3	2.95	2.94	-0.01	0.01	0.00	0.34
50	class3	2.79	2.71	-0.08	0.08	0.03	2.87
51	class3	2.7	2.77	0.07	0.07	0.03	2.59
52	class3	2.95	3.11	0.16	0.16	0.05	5.42
53	class4	3.35	3.34	-0.01	0.01	0.00	0.30
54	class4	3.25	3.61	0.36	0.36	0.11	11.08
55	class4	3.17	3.31	0.14	0.14	0.04	4.42
56	class4	3.39	3.36	-0.03	0.03	0.01	0.88

Table 4.43 Results of predicting pH values from Lab values using ANN(cont)

Mango No.	Class	Actual pH Value	Prediction pH Value	Error	Absolute error	Relative error	Percentage error
57	class4	3.37	3.55	0.18	0.18	0.05	5.34
58	class4	3.44	3.35	-0.09	0.09	0.03	2.62
59	class4	2.88	2.86	-0.02	0.02	0.01	0.69
60	class4	3.42	3.32	-0.10	0.1	0.03	2.92
61	class4	3.42	3.35	-0.07	0.07	0.02	2.05
62	class4	3.31	3.26	-0.05	0.05	0.02	1.51
63	class4	2.95	3.20	0.25	0.25	0.08	8.47
64	class4	3.43	3.39	-0.04	0.04	0.01	1.17
65	class4	3.28	3.28	0.00	0	0.00	0.00
66	class4	3.13	3.41	0.28	0.28	0.09	8.95
67	class4	3.4	3.34	-0.06	0.06	0.02	1.76
68	class4	3.05	3.02	-0.03	0.03	0.01	0.98
69	class4	2.85	3.05	0.20	0.2	0.07	7.02
70	class4	3.37	3.36	-0.01	0.01	0.00	0.30
71	class4	3.15	3.36	0.21	0.21	0.07	6.67
72	class4	4.14	4.01	-0.13	0.13	0.03	3.14
73	class5	4.23	4.40	0.17	0.17	0.04	4.02
74	class5	3.83	4.60	0.77	0.77	0.20	20.10
75	class5	4.15	4.18	0.03	0.03	0.01	0.72
76	class5	4.25	4.31	0.06	0.06	0.01	1.41
77	class5	4.13	4.38	0.25	0.25	0.06	6.05
78	class5	3.96	3.98	0.02	0.02	0.01	0.51
79	class5	3.82	4.16	0.34	0.34	0.09	8.90
80	class5	3.83	3.62	-0.21	0.21	0.05	5.48
81	class5	3.88	3.73	-0.15	0.15	0.04	3.87
82	class5	4.09	4.22	0.13	0.13	0.03	3.18
83	class5	3.95	4.24	0.29	0.29	0.07	7.34
84	class5	3.94	4.18	0.24	0.24	0.06	6.09
85	class5	3.94	4.11	0.17	0.17	0.04	4.31

Table 4.44 provides the analysis of absolute, relative, and percentage errors for pH predictions across six classes (class 0 to class 5). The average absolute errors ranged from 0.1050 (class 0) to 0.2177 (class 5), with an overall average of 0.1575. Standard deviations for absolute errors varied from 0.1029 (class 4) to 0.1921 (class 5), averaging 0.1471 across all classes. Relative errors ranged from 0.0351 (class 4) to 0.0909 (class 1), with an average of 0.0560. Standard deviations for relative errors ranged from 0.0331 (class 4) to 0.0789 (class 1), averaging 0.0552. Percentage errors varied from 3.5133%

(class 4) to 9.0946% (class 1), with an average of 5.6021%. Standard deviations for percentage errors ranged from 3.3088 (class 4) to 7.8903 (class 1), averaging 5.5171%. Overall, the predictions showed relatively low errors across most classes, indicating generally accurate pH predictions with minor variations.

Table 4.44 Absolute and relative error of pH predictions (ANN and Lab values)

Class	class0	class1	class2	class3	class4	class5	Average
Absolute error	0.1050	0.2044	0.2185	0.1311	0.1130	0.2177	0.1575
Standard deviation	0.1244	0.1591	0.1834	0.1078	0.1029	0.1921	0.1471
Relative error	0.0515	0.0909	0.0831	0.0457	0.0351	0.0554	0.0560
Standard deviation	0.0591	0.0789	0.0713	0.0374	0.0331	0.0504	0.0552
Percentage error	5.1455	9.0946	8.3064	4.5746	3.5133	5.5379	5.6021
Standard deviation	5.9105	7.8903	7.1335	3.7353	3.3088	5.0417	5.5171

Table 4.45 compares the average actual pH values and predicted pH values, along with their standard deviations across six classes (class 0 to class 5). The average actual pH values ranged from 2.0175 (class 0) to 4.0000 (class 5), with corresponding standard deviations ranging from 0.0625 (class 0) to 0.1524 (class 5). The predicted pH values ranged from 2.0692 (class 0) to 4.1514 (class 5), with varying standard deviations ranging from 0.1682 (class 1) to 0.3481 (class 2). Overall, the predicted pH values generally aligned closely with the actual pH values, although there were slight deviations across different classes, particularly notable in class 2, where the standard deviation of predicted pH values was higher.

Table 4.45 Average and standard deviation of actual and predicted pH values (ANN and Lab values)

	class0	class1	class2	class3	class4	class5
Average pH (Actual values)	2.0175	2.3111	2.6738	2.8628	3.2875	4.0000
Standard deviation	0.0625	0.2511	0.2380	0.1066	0.2747	0.1524
Average pH (Predicted values)	2.0692	2.3378	2.6569	2.8906	3.3011	4.1514
Standard deviation	0.1732	0.1682	0.3481	0.1456	0.1740	0.2581

Figure 4.12 shows the comparison between the pH values predicted by ANN and the actual values. The predicted mango Brix values were very good compared to the actual values.

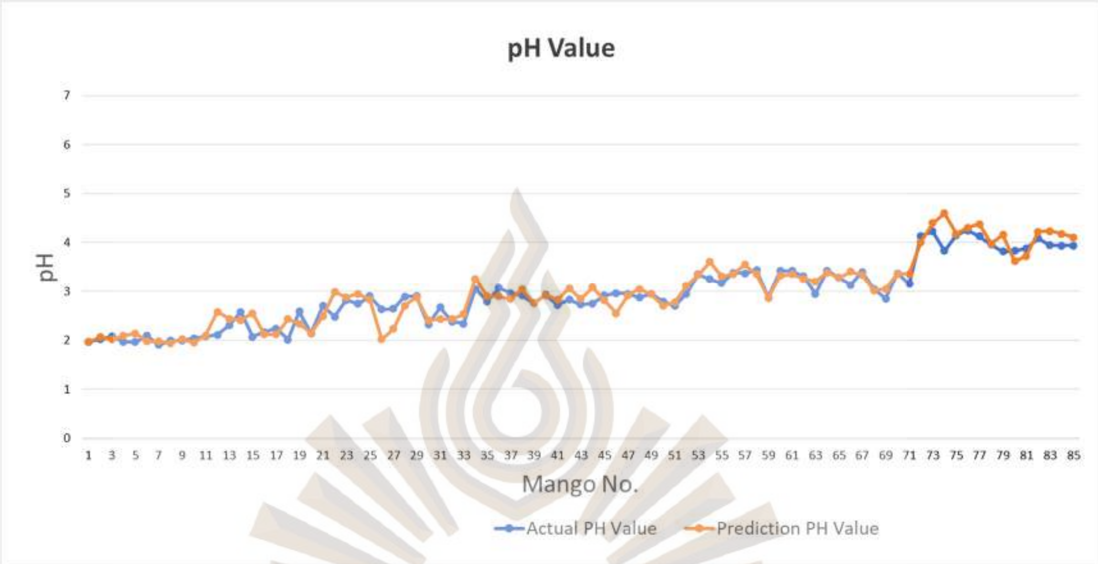

Figure 4.12Graphs shown the actual and predicted Brix values when predicting using ANN and Lab values.

Source: Researcher, 2024

Tables 4.40, 4.41, 4.42, 4.43, 4.4, and 4.45 above show the relationship between Lab values and pH values predicted by linear regression equations and ANN predictors. In terms of average error, ANN performed the best.

Tables 4.46 and 4.47 show the R^2 values between the actual and predicted values of all Brix and pH.

Table 4.46 R^2 of the methods for predicting Brix values

Method to predict Brix values	Features	R^2
Linear regression	Softmax	0.96
ANN	Softmax	0.96
Linear regression	RGB	0.89
ANN	RGB	0.94
Linear regression	Lab	0.9
ANN	Lab	0.94

Table 4.47 R^2 of the methods for predicting pH values

Method to predict pH values	Features	R^2
Linear regression	Softmax	0.91
ANN	Softmax	0.89
Linear regression	RGB	0.71
ANN	RGB	0.89
Linear regression	Lab	0.86
ANN	Lab	0.88

The best R^2 values for predicting Brix values came from ANN and linear regression when the softmax values from classifying the whole mango image with a VGG16 classifier were used. These values were 0.96. Using three RGB points and three Lab points, the same R^2 of 0.94 was achieved. When predicting pH values, linear regression could provide the highest R^2 of 0.91 when using the softmax values obtained from a VGG16 classifier, while the ANN could obtain a slightly lower R^2 of 0.89. The R^2 value was 0.89 based on three RGB points. The results revealed that the sofmax features were good for predicting mango Brix and pH values

Chapter 5

Conclusion and Recommendations

5.1 Conclusion

The methods for classifying mango ripeness were studied in this work. Four machine learning classifiers consisting of CNN, MobileNet, ResNet50, and VGG16 were examined to recognize six stages of mango ripeness. The results revealed that VGG16 could accurately classify mangoes at different ripeness stages. To enhance comprehension of the accuracy achievable with each method, the inclusion of additional images for training and testing machine learning classifiers is required. Furthermore, we should study the internal properties related to the external appearance of mangoes to determine their properties using image processing and machine learning.

This study aimed to investigate the relationships between the visual color data, softmax layer data, and the chemical properties (pH and Brix values) of mangoes at different maturity stages. We collected extensive data, including RGB and Lab color values, softmax layer outputs from convolutional neural networks, and pH and Brix measurements. The collected data were analyzed using two distinct methods: linear regression (LR) and ANN. We focused on six specific relationships: RGB and Brix, RGB and pH, Lab and Brix, Lab and pH, softmax and Brix, and softmax and pH. Our goal was to determine which method provided the most accurate predictions for each of these relationships.

Through a detailed analysis, for Brix prediction from softmax values, both LR and ANN achieved an R^2 of 0.96. For pH prediction from softmax values, LR achieved an R^2 of 0.91, while ANN achieved an R^2 of 0.89. For Brix prediction from RGB values, LR achieved an R^2 of 0.89, while ANN achieved an R^2 of 0.94. For pH prediction from RGB values, LR achieved an R^2 of 0.71, while ANN achieved an R^2 of 0.89. For Brix prediction from Lab values, LR achieved an R^2 of 0.9, while ANN achieved an R^2 of

0.94. For pH prediction from Lab values, LR achieved an R^2 of 0.86, while ANN achieved an R^2 of 0.88. Linear regression provided moderate prediction accuracy across all relationships, while ANN demonstrated superior performance in most cases. Specifically, ANN was particularly effective in predicting the pH values from softmax layer data, surpassing linear regression in prediction accuracy. In this scenario, the effectiveness of ANN highlights the potential of deep learning models for capturing complex patterns within the data.

Based on what we found, future research could focus on a number of areas, such as: improving and tweaking ANN models to make them more accurate; looking into hybrid models that combine the best features of linear regression and ANN; creating better feature extraction techniques to make the data we use more useful; and putting in place scalable and automated systems for judging mango maturity in farming that give accurate and real-time results. In conclusion, this study demonstrates the potential of machine learning methods, particularly ANN, in predicting the chemical properties of mangoes based on visual and deep learning-derived features. These findings pave the way for more advanced applications of machine learning in agriculture, offering promising solutions for quality control and maturity assessment of fruits

References

Aherwadi, N., Mittal, U., Singla, J., Jhanjhi, N. Z., Yassine, A., & Hossain, M. S. (2022). Prediction of Fruit Maturity, Quality, and Its Life Using Deep Learning Algorithms. *Electronics*, *11*(24). https://doi.org/10.3390/electronics11244100

Alqahtani, O. M. (2024). *Efficient Deep Learning for Plant Disease Classification in Resource Constrained Environment* (Doctoral dissertation, University of Georgia). Retrieved from https://www.proquest.com/dissertations-theses/efficient-deep-learning-plant-disease/docview/3065560830/se-2

Bouguezzi, S., Fredj, H. B., Belabed, T., Valderrama, C., Faiedh, H., & Souani, C. (2021). An Efficient FPGA-Based Convolutional Neural Network for Classification:Ad-MobileNet. *Electronics*,*10*(18). https://doi.org/10.3390/electronics10182272

Bureau, S. (2009). The use of non-destructive methods to analyse fruit quality. *Fresh produce*, *3*(1), 23-34.

Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges, and trends. *Neurocomputing*, *408*,189–215. https://doi.org/10.1016/j.neucom.2020.04.084.

Chomchalow, N., & Songkhla, P. N. (2008). Thai mango export: a slow-but-sustainable development. *Assumption University Journal of Technology*, *12*(1), 1-8.

Ciccullo, F., Cagliano, R., Bartezzaghi, G., & Perego, A. (2021). Implementing the circular economy paradigm in the agri-food supply chain: The role of food waste prevention technologies. *Resources, Conservation and Recycling*, *164*. https://doi.org/10.1016/j.resconrec.2020.105114

Çınar, A., Yıldırım, M., & Eroğlu, Y. (2021). Classification of Pneumonia Cell Images Using Improved ResNet50 Model. *Traitement du Signal*, *38*(1), 165-173. https://doi.org/10.18280/ts.380117

References (Cont.)

El-Ramady, H. R., Domokos-Szabolcsy, É., Abdalla, N. A., Taha, H. S., & Fári, M. (2015). Postharvest Management of Fruits and Vegetables Storage. In *Sustainable Agriculture Reviews* (pp. 65-152). https://doi.org/10.1007/978-3-319-09132-7_2

Elngar, A. A., Arafa, M., Fathy, A., Moustafa, B., Mahmoud, O., Shaban, M., … Fawzy, N. (2021). Image classification based on CNN: a survey. *Journal of Cybersecurity and Information Management*, *6*(1), 18-50.

Eyarkai Nambi, V., Thangavel, K., Manickavasagan, A., & Shahir, S. (2017). Comprehensive ripeness-index for prediction of ripening level in mangoes by multivariate modelling of ripening behaviour. *International Agrophysics*, *31*(1), 35-44. https://doi.org/10.1515/intag-2016-0025

Lu, Y., Huang, Y., & Lu, R. (2017). Innovative Hyperspectral Imaging-Based Techniques for Quality Evaluation of Fruits and Vegetables: A Review. *Applied Sciences*, *7*(2). https://doi.org/10.3390/app7020189

Magwaza, L. S., Opara, U. L., Nieuwoudt, H., Cronje, P. J. R., Saeys, W., & Nicolaï, B. (2011). NIR Spectroscopy Applications for Internal and External Quality Analysis of Citrus Fruit—A Review. *Food and Bioprocess Technology*, *5*(2), 425-444. https://doi.org/10.1007/s11947-011-0697-1

Mazen, F. M. A., & Nashat, A. A. (2019). Ripeness Classification of Bananas Using an Artificial Neural Network. *Arabian Journal for Science and Engineering*, *44*(8), 6901-6910. https://doi.org/10.1007/s13369-018-03695-5

Mim, F. S., Galib, S. M., Hasan, M. F., & Jerin, S. A. (2018). Automatic detection of mango ripening stages – An application of information technology to botany. *Scientia Horticulturae*,*237*,156-163.

Mimma, N.-E. A., Ahmed, S., Rahman, T., Khan, R., & Tang, Z. (2022). Fruits Classification and Detection Application Using Deep Learning. *Scientific Programming*, *2022*, 1-16. https://doi.org/10.1155/2022/4194874

References (Cont.)

Miriti, E. (2016). Classification of selected apple fruit varieties using Naive Bayes (Doctoral dissertation). Retrieved from https://erepository.uonbi.ac.ke/handle/11295/97285

Mkhathini, K. M. (2014). *The influence of physiological maturity on physicochemical quality attributes, value addition and storage life of three processed products of peach 'Landarces' produced at Impendle, South Africa* (Literature review). Retrieved from https://efwe.ukzn.ac.za/wp-content/uploads/2021/11/Mkhathini_KM.pdf

Nan, Y., Ju, J., Hua, Q., Zhang, H., & Wang, B. (2022). A-MobileNet: An approach of facial expression recognition. *Alexandria Engineering Journal*, *61*(6), 4435-4444. https://doi.org/10.1016/j.aej.2021.09.066

Ngugi, L. C., Abelwahab, M., & Abo-Zahhad, M. (2021). Recent advances in image processing techniques for automated leaf pest and disease recognition–A review. *Information processing in agriculture*, *8*(1), 27-51.

Nguyen, C., Sagan, V., Maimaitiyiming, M., Maimaitijiang, M., Bhadra, S., & Kwasniewski, M. T. (2021). Early Detection of Plant Viral Disease Using Hyperspectral Imaging and Deep Learning. *Sensors (Basel)*, *21*(3). https://doi.org/10.3390/s21030742

Osisanwo F.Y., A. J. E. T., Awodele O., Hinmikaiye J. O., Olakanmi O.&Akinjobi J. (2017). Supervised Machine Learning Algorithms: Classification and Comparison. *International Journal of Computer Trends and Technology*, *48*(3), 128-138. https://doi.org/10.14445/22312803/ijctt-v48p126

Prabhu, A., & Mamatha, C. (2022, July). Computer Vision System in Identifying the Ripening Stages of Mango–Alphonso Cultivar. In *2022 IEEE International Conference on Data Science and Information System (ICDSIS)* (pp. 1-6). IEEE.

References (Cont.)

Rezende, E., Ruppert, G., Carvalho, T., Theophilo, A., Ramos, F., & Geus, P. D. (2018). Malicious software classification using VGG16 deep neural network's bottleneck features. In *Information Technology-New Generations: 15th International Conference on Information Technology* (pp. 51-59). Springer International Publishing.

Rizzo, M., Marcuzzo, M., Zangari, A., Gasparetto, A., & Albarelli, A. (2023). Fruit ripeness classification: A survey. *Artificial Intelligence in Agriculture*, *7*, 44-57. https://doi.org/10.1016/j.aiia.2023.02.004

Seymour, G. B., Østergaard, L., Chapman, N. H., Knapp, S., & Martin, C. (2013). Fruit development and ripening. *Annual review of plant biology*, *64*(1), 219-241.

Singh, A., Vaidya, G., Jagota, V., Darko, D. A., Agarwal, R. K., Debnath, S., & Potrich, E. (2022). Recent advancement in postharvest loss mitigation and quality management of fruits and vegetables using machine learning frameworks. *Journal of Food Quality*, *2022*(1), 6447282.

Vidanapathirana, R. (2019). Safety issues in fresh fruit and vegetable supply chains in Sri Lanka: A review. *Proceedings of the 1st International Conference on Agriculture, Food Security and Safety*, *1*, 17.

Wu, J. (2017). Introduction to convolutional neural networks. *National Key Lab for Novel Software Technology. Nanjing University. China*, *5*(23), 495.

Yang, B., & Xu, Y. (2021). Applications of deep-learning approaches in horticultural research: a review. *Hortic Res*, *8*(1), 123. https://doi.org/10.1038/s41438-021-00560-9

Yang, R., Lu, X., Huang, J., Zhou, J., Jiao, J., Liu, Y., … Gu, P. (2021). A multi-source data fusion decision-making method for disease and pest detection of grape foliage based on ShuffleNet V2. *Remote Sensing, 13*(24). https://doi.org/10.3390/rs13245102

Zhena, O. P., Hashima, N., & Maringgala, B. (2020). Quality evaluation of mango using non-destructive approaches: A review. *Journal of Agricultural and Food Engineering*, *1*(0003).

Biography

Name	Xiangchen Dong
Date of birth	11 February 2000
Place of birth	Hebei/China
Education background	Beijing University of Chemical Technology, Bachelor of Communication Engineering, 2021
Address	Transportation Management Station Family Courtyard No. 005, Xiong county, Xiong'an New Area, Hebei Province, China
Email Address	dongc0211@gmail.com